

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Run on: April 21, 2003, 19:25:40 ; Search time 5085.18 Seconds

Title: US-10-003-690-1

Perfect score: 2980  
Sequence: 1 gtcgaccacgcgtccgcg.....aaaaaaaaagggcgccgcta 2980

Scoring table:	IDENTITY_NUC	Canext 1 0
Canext 10 0		

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

```

1:  gb_ba:*
2:  gb_hcg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_srs:*
12: gb_sy:*
13: gb_un:*
14: gb_v1:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_om:*
21: em_ov:*
22: em_or:*
23: em_pdt:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_srs:*
28: em_un:*
29: em_v1:*
30: em_hcg_hum:*
31: em_hcg_hum:*
32: em_hcg_fav:*
33: em_hcg_other:*
34: em_hcg_mus:*
35: em_hcg_pdl:*
36: em_hcg_rin:*
37: em_hcg_vrt:*
38: em_hcg_vrt:*
39: em_hcg_hum:*
40: em_hcg_omus:*
41: em_hcg_other:*

```

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB	ID	Description
1	2868	96.2	3109	9	AF479827	AF479827 Homo sapi
2	2734	92.4	3007	9	AF479826	AF479826 Homo sapi
3	2708	90.9	2720	9	HSMB05307	ALB34275 Homo sapi
4	2576	86.4	2576	9	AB058714	AB058714 Homo sapi
5	2259	75.8	2188	9	AX166526	AX166526 Sequence
6	1953.6	65.6	2128	9	BC016681	BC016681 Homo sapi
7	969	32.5	1014	9	HSMB03233	ALB31945 Homo sapi
8	961	32.2	2025	6	AX327995	AX327995 Sequence
9	961	32.2	2219	6	AX327993	AX327993 Sequence
10	818	27.4	1956	9	HSAB6701	AJ006701 Homo sapi
11	678.8	22.8	134792	9	AC020922	AC020922 Homo sapi
12	491.8	16.5	3156	9	AK074411	AK074411 Homo sapi
13	446	15.0	3933	3	AB014885	AB014885 Halocynthia
14	391.2	13.0	3053	3	AF316542	AF316542 Caenorhab
15	373.6	12.5	5609	3	AT060288	AT060288 Drosophila
16	371	12.4	38000	9	AC0089374	AC0089374 Homo sapi
17	366.4	12.3	226600	2	AC079583	AC079583 Mus muscu
18	314	10.5	210105	2	AC125948	AC125948 Rattus no
19	307.2	10.3	1873	9	BC024291	BC024291 Homo sapi
20	296.8	10.0	2333	9	AF020089	AF020089 Homo sapi
21	269	9.0	3529	9	AB088047	AB088047 Homo sapi
22	269	9.0	3469	9	AB049127	AB049127 Homo sapi
23	267.4	9.0	2462	6	AX399887	AX399887 Sequence
24	267.4	9.0	3326	9	AY057448	AY057448 Homo sapi
25	267.4	9.0	3312	6	AX305105	AX305105 Sequence
26	267.4	9.0	3392	6	AX305106	AX305106 Sequence
27	267.4	9.0	4917	6	AB058763	AB058763 Homo sapi
28	261.6	8.8	1594	6	AX056390	AX056390 Sequence
29	259.4	8.7	3170	6	AX305103	AX305103 Sequence
30	259.4	8.7	3250	6	AX305104	AX305104 Sequence
31	245	8.2	2650	10	AB020480	AB020480 Rattus no
32	245	8.2	4092	10	AF106937	AF106937 Rattus no
33	240.6	8.1	2112	6	AX375128	AX375128 Sequence
34	240.6	8.1	2112	9	AF387638	AF387638 Homo sapi
35	240.6	8.1	2222	6	AX375131	AX375131 Sequence
36	240.6	8.1	2701	9	BC008771	BC008771 Homo sapi
37	240.6	8.1	2946	9	HSSSPKMK	X97630 H.sapiens m
38	238.4	8.0	2572	10	RRMAR2	Z83869 R.norvegicu
39	237.2	8.0	1749	9	AF020310	AF020310 Drosophil
40	228.8	7.7	2763	10	MMKEM	X70764 M.musculus
41	227.8	7.6	2681	3	AF181649	AF181649 Drosophil
42	225.4	7.6	1869	8	ATSRK12	X94757 A.thaliana
43	225.4	7.5	4498	10	MMU11494	AY093170 Arabidops
44	224.2	7.4	2352	6	AX262512	U11494 Mus musculu
45	221.4	7.4	2352	6	AX262512	AX262512 Sequence

FEATURES  
source  
Human Genes Research, Institute of Genetics, Fudan University, 220  
Handan Road, Shanghai 200433, P. R. China  
location/Qualifiers  
1..3109  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="19"  
278..2614  
CDS

BASE COUNT	611 a	1033 c	957 g	508 t
ORIGIN				

Query Match	96.28;	Score 2868;	DB 9;	Length 3109;
Best Local Similarity	99.38;	Pred. No. 0;		
Matches 2880; Conservative	0;	Mismatches 20;	Indels 0;	Gaps 0

QY	18	CGACGCGTGGGGCGGACGCGTGGGGGAGAACGCTGGGGGGGAGACGCGTGGGGCGGGGAGACA	77
Db	210	CGGAGAGACGGGGCGACGGCCCGCAGGGGGGGCGCGCCGGGGAGACCGGTGGGGCGGGGAGCCA	269
QY	78	AGGCGACCATGTCTGTCGGGGGCGAAAGAGAGAGTGGGGGCTCTCCCGCTACCACTCC	137
Db	270	AGGGACCATGTCTGTCGGGGGCGCAAGAGAGAGTGGGGGCTCTCCCGCTACCACTCC	329
QY	138	CCGACCCCGACCCCGCACCCCGACGACGCCCAATATGTGGGCCCTATGGCTGGAGA	197
Db	330	CCGACCCCGACCCCGCACCCCGACGACGCCCAATATGTGGGCCCTATGGCTGGAGA	389
QY	198	AGACGCTGGGCGAAAGACAGACAGGGCTGGTTAACTCGGGGTCCACGACATCAAGGGTTC	257
Db	390	AGACGCTGGGCGAAAGACAGACAGGGCTGGTTAACTCGGGGTCCACGACATCAAGGGTTC	449
QY	258	AGAGGTGCCATCAAGATCGTGAACCGGGAGAACGCTCGGAGTGGTGTGATGAAGG	317
Db	450	AGAGGTGCCATCAAGATCGTGAACCGGGAGAACGCTCGGAGTGGTGTGATGAAGG	509
QY	318	TGGAGCGGAGATCGCCATCTCTGAAGCTCATCGAACCACCAATGTCTCAAGCTCCAGC	377
Db	510	TGGAGCGGAGATCGCCATCTCTGAAGCTCATCGAACCACCAATGTCTCTCAAGCTCCAGC	569
QY	378	ACGCTACGAGAACAGAAATATTTGTACCTGTTCTGGACACAGTTCGGGGGGGTAGC	437
Db	570	ACGCTACGAGAACAGAAATATTTGTACCTGTTCTGGACACAGTTCGGGGGGGTAGC	629
QY	438	TATTTCGACTACCTGGTAAAGAGGGGAGACTGACGCCCAAGGAGGCCCGAAATCTTCC	497
Db	630	TATTTCGACTACCTGGTAAAGAGGGGAGACTGACGCCCAAGGAGGCCCGAAATCTTCC	689
QY	498	GCGAGATTTGTCTTCGGCTGGACTTTCGCCACAGCTACTCCTCATCTGCGACAGAGACTAA	557
Db	690	GCGAGATTTGTCTTCGGCTGGACTTTCGCCACAGCTACTCCTCATCTGCGACAGAGACTAA	749
QY	558	AGCCCGAGAACCTCTTTTGGATGAGAAAAACAATCCGATTTGGACAGACTTGGCGATGG	617
Db	750	AGCCCGAGAACCTCTTTTGGATGAGAAAAACAATCCGATTTGGACAGACTTGGCGATGG	809

QY	618	CGTCCCTCAGGTGGGGGACACCTTCGTGGACACCAAGCTGCGGTGCCCATTTAGCGT	677
Db	810	CGTCCCTCAGGTGGGGGACACCTTCGTGGACACCAAGCTGCGGTGCCCATTTAGCGT	869
QY	678	GTCCAGAGGTATTAAGGGGGAAAAATATGATGGCCCGCGGGGACAGCATGTGGAGCTGTG	737
Db	870	GTCCAGAGGTATTAAGGGGGAAAAATATGATGGCCCGCGGGGACAGCATGTGGAGCTGTG	929
QY	738	GAGTCATCTCTTGCCCTCGCTCGGTGGGGGCTCTGCGCTTTATGACGACAACTCCGCC	797
Db	930	GAGTCATCTCTTGCCCTCGCTCGGTGGGGGCTCTGCGCTTTATGACGACAACTCCGCC	989
QY	798	AGCTGTGGAGAAAGGTGMAAGGGGGCGTCTTCACATGCGCCCACTTCATTCCTCCAGAT	857
Db	990	AGCTGTGGAGAAAGGTGMAAGGGGGCGTCTTCACATGCGCCCACTTCATTCCTCCAGAT	104
QY	858	GCCGAGCCCTCTCAGGGGAATGATCGAATGTGAGCCCGGAAAAAAGGCTCACTGTGAGC	917
Db	1050	GCCGAGCCCTCTCAGGGGAATGATCGAATGTGAGCCCGGAAAAAAGGCTCACTGTGAGC	1100
QY	918	AAATTCAGAAACATCTCTTGTAACCTAGACGGGAAACACGAGCCACACCTGCTCGAGC	977
Db	1110	AAATTCAGAAACATCTCTTGTAACCTAGACGGGAAACACGAGCCACACCTGCTCGAGC	1166
QY	978	CAGCCCTTGCGCCCGGGTAGCCATGCGGAGCTTGCCATCCAGCGAGCTGAGACCCG	1037
Db	1170	CAGCCCTTGCGCCCGGGTAGCCATGCGGAGCTTGCCATCCAGCGAGCTGAGACCCG	1222
QY	1038	ACGTCCTAGAGAGCATGGCATCAGTGGGCTGCTTACGAGGACCGGAGAGCTGCATCGG	1097
Db	1230	ACGTCCTAGAGAGCATGGCATCAGTGGGCTGCTTACGAGGACCGGAGAGCTGCATCGG	1288
QY	1098	AGCTGCGCAGTGAAGAGAGAAACCAAGAAAGATGATATATTATCTGTTTGTGATCGGA	1157
Db	1290	AGCTGCGCAGTGAAGAGAGAAACCAAGAAAGATGATATATTATCTGTTTGTGATCGGA	1344
QY	1158	AGGAGGGGTATCCAGCTGTGAGGACCAAGACCTGCTCCCGGATGATGTTGACCCG	1217
Db	1350	AGGAGGGGTATCCAGCTGTGAGGACCAAGACCTGCTCCCGGATGATGTTGACCCG	1409
QY	1218	CCCGGAAGCGTGTGATTTCCATGCTGTAAGCCCTCACGGGAACGGGAGCACAGAGCGGA	1277
Db	1410	CCCGGAAGCGTGTGATTTCCATGCTGTAAGCCCTCACGGGAACGGGAGCACAGAGCGGA	1466
QY	1278	AGTCACATGGAAGTCGTGAGATCACCGATGCGGGGTGTGCTCTCCCTGTACCACCC	1337
Db	1470	AGTCACATGGAAGTCGTGAGATCACCGATGCGGGGTGTGCTCTCCCTGTACCACCC	1529
QY	1338	GACGGGCTTGTGATGTGGCCACACACAGCCAGAGATCCCGTAGGTGATAGTGAAGCTTCA	1397
Db	1530	GACGGGCTTGTGATGTGGCCACACACAGCCAGAGATCCCGTAGGTGATAGTGAAGCTTCA	1589
QY	1398	CGGGTGTGTCTTCAGCCCTCTAAGCAGGCCCAAGAGATCCGGTCTTTCTTTTACCGG	1457
Db	1590	CGGGTGTGTCTTCAGCCCTCTAAGCAGGCCCAAGAGATCCGGTCTTTCTTTTACCGG	1649
QY	1458	AGCCGGGGCTGTGATGTAGAGCTCGAGGGGGGGGCTCCCGACTTTCCAAAAAGCAGAGC	1517
Db	1650	AGCCGGGGCTGTGATGTAGAGCTCGAGGGGGGGGCTCCCGACTTTCCAAAAAGCAGAGC	1709
QY	1518	TGCTTCTTGCGGGCCCGCAGGGGTGTGGGGCGCCGAGGAGCAGCCCTCGCCCGAGTAGC	1577
Db	1710	TGCTTCTTGCGGGCCCGCAGGGGTGTGGGGCGCCGAGGAGCAGCCCTCGCCCGAGTAGC	1769
QY	1578	GCTCACAACCCGAGCCCGGCGCCCAAGGCTCCCGGCTCTCTTGAGCGGACCCCTTGC	1637
Db	1770	GCTCACAACCCGAGCCCGGCGCCCAAGGCTCCCGGCTCTCTTGAGCGGACCCCTTGC	1829
QY	1638	ACTGCGCTCTGCAACGCCCCGGGGCAGTCTCACCGGAGACCCGGGGACAACACACCC	1697
Db	1830	ACTGCGCTCTGCAACGCCCCGGGGCAGTCTCACCGGAGACCCGGGGACAACACACCC	1889
QY	1698	CCAGCCCCGGGCTGCGGGGAGACCGCTGTGAGAGATGCTTCAACTCCATCCGA	1757

```

1890 CCAGCCCCGGGCGTGGCGGGGAGCCCGCTGAGAGAGTCTGCTCAACCTCCATCCGCA 1949
1758 AAGAGCTCTGGGCTCCCTCCGCTTCAACCGGGGCAAGATGAGGCTCCATACGCTGAG 1817
1950 AAGAGCTCTGGGCTCCCTCCGCTTCAACCGGGGCAAGATGAGGCTCCATACGCTGAG 2009
1818 AAGATGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1877
2010 AAGATGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2069
1878 ACTTCATCTCTGCTGAGCAAGAAAGAAATATCTCTGCTGCTAAGAGAGAGAGAGAGAG 1937
2070 ACTTCATCTCTGCTGAGCAAGAAAGAAATATCTCTGCTGCTAAGAGAGAGAGAGAGAG 2129
1938 GAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1997
2130 GAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2189
1998 TGTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2057
2190 TGTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2249
2058 AAAAGCCGCTCCGCTTCCAGAGTGAATCAAGCTCTCTGAGGGTCCAGAGCCCTCCGCG 2117
2250 AAAAGCCGCTCCGCTTCCAGAGTGAATCAAGCTCTCTGAGGGTCCAGAGCCCTCCGCG 2309
2118 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2177
2310 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2369
2178 CCAGCGCTGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2237
2370 CCAGCGCTGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2429
2238 ACCAGCCCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2297
2430 ACCAGCCCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2489
2298 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2357
2490 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2549
2358 CTCGCCGCTCCGCTTCCAGAGTGAATCAAGCTCTCTGAGGGTCCAGAGCCCTCCGCG 2417
2550 CTCGCCGCTCCGCTTCCAGAGTGAATCAAGCTCTCTGAGGGTCCAGAGCCCTCCGCG 2609
2418 CTCGCCGCTCCGCTTCCAGAGTGAATCAAGCTCTCTGAGGGTCCAGAGCCCTCCGCG 2477
2610 CTCGCCGCTCCGCTTCCAGAGTGAATCAAGCTCTCTGAGGGTCCAGAGCCCTCCGCG 2669
2478 AACTGATCAATCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2537
2670 AACTGATCAATCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2729
2538 CGGCGCTTCCGCTTCCAGAGTGAATCAAGCTCTCTGAGGGTCCAGAGCCCTCCGCG 2597
2730 CGGCGCTTCCGCTTCCAGAGTGAATCAAGCTCTCTGAGGGTCCAGAGCCCTCCGCG 2789
2598 GGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2657
2790 GGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2849
2658 AATGATCAATCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2717
2850 AATGATCAATCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2909
2718 AAGCAGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2777
2910 AAGCAGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2969
2778 CTCCTCCGCTTCCGCTTCCAGAGTGAATCAAGCTCTCTGAGGGTCCAGAGCCCTCCGCG 2837

```

```

Db 2970 CTCCTCCCTGCTCTCCCGCCAGACGCTTCTGATGAGATTTCTCTCCGGAAGAAATTC 3029
Qy 2838 TGGTTTCGGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2897
Db 3030 TGGTTTCGGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3089
Qy 2898 AAAAAAAAAAAAAAAAAA 2917
Db 3090 AAAAAAAAAAAAAAAAAA 3109

RESULT 2
AF479826 3007 bp mRNA linear PRI 12-MAR-2002
LOCUS Homo sapiens putative serine/threonine protein kinase mRNA,
DEFINITION complete cds.
ACCESSION AF479826
VERSION AF479826.1 GI:19401870
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 3007)
AUTHORS She,X.Y., Guo,J.H. and Yu,L.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-2002) School of Life Sciences, Laboratory of
Human Genes Research, Institute of Genetics, Fudan University, 220
Handan Road, Shanghai 200433, P. R. China
FEATURES
Source Location/Qualifiers
1..3007
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13"
/tissue.type="brain"
128..2512
/codon_start=1
/product="putative serine/threonine protein kinase"
/protein_id="FAL87697.1"
/db_xref="GI:19401871"
/translation="MVAGLLTGKSPSPDGVSPERKDEAVAGGSGEEDAEERGRHA
OYVPRLEKTKGQGTGLVKKVCHVCTGQKVAIKIVNREKLSVLMKEKRETAIIK
LIEHPVILKLDVYENKKYIYLVEHVSSELIDYLVKKRILPKEARKPRQIVSL
DFCHSYSTICHRDLPENLIDENNNIRIADFGMSLDVGSLSLETSGSPHAPCEYI
KKEKYDRADWMSGVILFALLVGLAFDDDNILROLLEVKRQVFMPIIPDCOS
LIRGATIEVEPERKLSLEOIKHPWYLGKEHPDCELPARGRVAMSLPSNGELDP
VLESMASLIGCFRDLRELRSEENOEKMIYLLIDRKERYSCEDDILPPRNDV
PPRKRVDSMWLSHRGRPRPKRMEVSLTDAGGGGSPVPTRALEMAHOHSORRSVS
GASTGISSEPLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
PPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
SRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLSIRNLS
FLVLDKPLSLSLKADIVHAFSLISLSHVSLSQTSFAEYKASGSPVFOKPVRFQYD
ISSBEPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
LADKNGAOTRRPAGAPRSLQPPGRDPPELSSPRNGPPPKKLLATNGPLP."

BASE COUNT 636 a 940 c 906 g 525 t
ORIGIN
Query Match 92.4%; Score 2754; DB 9; Length 3007;
Best local Similarity 100.0%; Pred. No. 0;
Matches 2754; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 164 CAGGCCCAATATGTGGGCCCCCTATCGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 223
Db 254 CAGGCCCAATATGTGGGCCCCCTATCGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 313
Qy 224 CTGGTTAACTCGGGGTCACATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 283
Db 314 CTGGTTAACTCGGGGTCACATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 373
Qy 284 CGGAGAGAGTGTGAGAGTGTGATGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 343
Db 374 CGGAGAGAGTGTGAGAGTGTGATGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 433

```

OY 344 CTCATGAACACCCACATGCTCTCAAGCTCCACGACGCTACGAGAACAAATATTG 403  
 DB 434 CTCATGAACACCCACATGCTCTCAAGCTCCACGACGCTACGAGAACAAATATTG 493  
 OY 404 TACCTGTTCTGGAGACGCTCGGGGGGTGAGTATTGACCTCTGTTAAAGAGGG 463  
 DB 494 TACCTGTTCTGGAGACGCTCGGGGGGTGAGTATTGACCTCTGTTAAAGAGGG 553  
 OY 464 AGACTGACCCCAAGAGAGGCGCCAAAGTTCTTCGCGCAGATTGTGTGCGCTGACCTC 523  
 DB 554 AGACTGACCCCAAGAGAGGCGCCAAAGTTCTTCGCGCAGATTGTGTGCGCTGACCTC 613  
 OY 524 TCCACAGCTACTCTCAATCTGCGACAGAGACCTAAAGCCGAGAACTGTTTGGATAG 583  
 DB 614 TCCACAGCTACTCTCAATCTGCGACAGAGACCTAAAGCCGAGAACTGTTTGGATAG 673  
 OY 584 AAAAACACATCGGACGCTGAGACTTCGGCATGCGGCTGCGCTGAGGTGGGGAGACCTC 643  
 DB 674 AAAAACACATCGGACGCTGAGACTTCGGCATGCGGCTGCGCTGAGGTGGGGAGACCTC 733  
 OY 644 CTGAGACACGCTGCGGCTCCCGCATTTATGCTGTCTCAGAGGTATTAAAGGGGAAAA 703  
 DB 734 CTGAGACACGCTGCGGCTCCCGCATTTATGCTGTCTCAGAGGTATTAAAGGGGAAAA 793  
 OY 704 TATGATGCGCGCGGCGACACATGTGAGCTGTGAGTCACTCTTCCCTGCTGCTG 763  
 DB 794 TATGATGCGCGCGGCGACACATGTGAGCTGTGAGTCACTCTTCCCTGCTGCTG 853  
 OY 764 GGGGCTGCTCCCTTGTGATGACGACACACTCCGCGAGCTGTGAGAAAGGAAAGGGGG 823  
 DB 854 GGGGCTGCTCCCTTGTGATGACGACACACTCCGCGAGCTGTGAGAAAGGAAAGGGGG 913  
 OY 824 GTCTTCCACATGCCCCACCTTCTCCAGATTGCGAGACCTCTGAGGGGAAATGATC 883  
 DB 914 GTCTTCCACATGCCCCACCTTCTCCAGATTGCGAGACCTCTGAGGGGAAATGATC 973  
 OY 884 GAAGTGGAGCCCGAAAAAGGCTCAAGTCTGGAGCAATTTAGAAACATCTTGTGACTA 943  
 DB 974 GAAGTGGAGCCCGAAAAAGGCTCAAGTCTGGAGCAATTTAGAAACATCTTGTGACTA 1033  
 OY 944 GGGGGAAAAACAGACGACGACGCTGAGACGACCTGGCGCGGGGTAGGACATG 1003  
 DB 1034 GGGGGAAAAACAGACGACGACGCTGAGACGACCTGGCGCGGGGTAGGACATG 1093  
 OY 1004 CGGAGCCTGCTCCACAGGAGAGCTGAGACCCCGAGCTCTTGAAGACATGCGATCACTG 1063  
 DB 1094 CGGAGCCTGCTCCACAGGAGAGCTGAGACCCCGAGCTCTTGAAGACATGCGATCACTG 1153  
 OY 1064 GGGCTGCTCAGGAGACGCGAGAGGCTGATCGAGCTGCGAGTGAAGAGGAGAACCA 1123  
 DB 1154 GGGCTGCTCAGGAGACGCGAGAGGCTGATCGAGCTGCGAGTGAAGAGGAGAACCA 1213  
 OY 1124 GAAAGATGATATATATCTGCTTTTGGATCGGAAGAGAGCGGTATCCAGCTGTGAGAC 1183  
 DB 1214 GAAAGATGATATATATCTGCTTTTGGATCGGAAGAGAGCGGTATCCAGCTGTGAGAC 1273  
 OY 1184 CAGAGCCTGCTCCCGGAAATGATGTGAGACCCCGGAAAGCTGTGATTTCCCATG 1243  
 DB 1274 CAGAGCCTGCTCCCGGAAATGATGTGAGACCCCGGAAAGCTGTGATTTCCCATG 1333  
 OY 1244 CTGAGCCTGACGAGGAAAGCGGACGACAGAGCGGAAGTCCATGGAATCTGAGCATCAC 1303  
 DB 1334 CTGAGCCTGACGAGGAAAGCGGACGACGAGCGGAAGTCCATGGAATCTGAGCATCAC 1393  
 OY 1304 GATCCGCGGGGTGTGCTCCCTGTATCCACCGGAGCGCTTGAAGTGGGCCACAC 1363  
 DB 1394 GATCCGCGGGGTGTGCTCCCTGTATCCACCGGAGCGCTTGAAGTGGGCCACAC 1453  
 OY 1364 AGCCAGAGATCCCGTAGGTGAGTGGAGGCTCTCAAGGCTCTCTCCAGCCCTCTAAGC 1423  
 DB 1454 AGCCAGAGATCCCGTAGGTGAGTGGAGGCTCTCAAGGCTCTCTCCAGCCCTCTAAGC 1513

---

OY 1424 AGCCAGAGATCCCGTCTTTCTTTCTTTTACCGGAGCGGGGGCTGGAGTAGGCTCGA 1483  
 DB 1514 AGCCAGAGATCCCGTCTTTCTTTCTTTTACCGGAGCGGGGGCTGGAGTAGGCTCGA 1573  
 OY 1484 GGGGGGGCTCCCGACTTCCAAAACGACAGACGCTGCTTCTGCGGGGCCAGGGGTGG 1543  
 DB 1574 GGGGGGGCTCCCGACTTCCAAAACGACAGACGCTGCTTCTGCGGGGCCAGGGGTGG 1633  
 OY 1544 GGGGGGGGAGACAGCGCCCGCCCGCCAGTGGCCCTCCACACCCCTGCGGGGCCCA 1603  
 DB 1634 GGGGGGGGAGACAGCGCCCGCCCGCCAGTGGCCCTCCACACCCCTGCGGGGCCCA 1693  
 OY 1604 GGGCTCCCGGCTCTCTGCGGGGACCCCTTGTGACCTGCTGTGACAGAGCCCGGGGC 1663  
 DB 1694 GGGCTCCCGGCTCTCTGCGGGGACCCCTTGTGACCTGCTGTGACAGAGCCCGGGGC 1753  
 OY 1664 AGTCCACCGGGAGACCCCGGGGACACACACACCCCGGCGGTGCGGGGGA 1723  
 DB 1754 AGTCCACCGGGAGACCCCGGGGACACACACACCCCGGCGGTGCGGGGGA 1813  
 OY 1724 GCGGCTGGAGAGTGTCTCAACTCCATCCGCAACAGCTTCTGCGGCTCTCCCTT 1783  
 DB 1814 GCGGCTGGAGAGTGTCTCAACTCCATCCGCAACAGCTTCTGCGGCTCTCCCTT 1873  
 OY 1784 CACCGGCGCAAGATCAGGTCTCCCTACCGCTGAGAGATGTCCAGTTGACGCGAGATCC 1843  
 DB 1874 CACCGGCGCAAGATCAGGTCTCCCTACCGCTGAGAGATGTCCAGTTGACGCGAGATCC 1933  
 OY 1844 TCCCGGAGCTGGCAAAAGCTCTGTGTTGGGAACCTCATCTCTTGAACAAAGAA 1903  
 DB 1934 TCCCGGAGCTGGCAAAAGCTCTGTGTTGGGAACCTCATCTCTTGAACAAAGAA 1993  
 OY 1904 CAAATATCTCTGCTCTTAAAGAGCAAACTCTCAGAGATCAAAACACATCTCTCAT 1963  
 DB 1994 CAAATATCTCTGCTCTTAAAGAGCAAACTCTCAGAGATCAAAACACATCTCTCAT 2053  
 OY 1964 GCTTTCTGTGATCCCGAGCTGAGTCAAGTGTGCTGACAGACAGCTTCAAGGGC 2023  
 DB 2054 GCTTTCTGTGATCCCGAGCTGAGTCAAGTGTGCTGACAGACAGCTTCAAGGGC 2113  
 OY 2024 GAGTACAGAGGCGAGTGGGGGCGCTCCCTCTTCCAAAAGCCCGCTTCCAGGTGAC 2083  
 DB 2114 GAGTACAGAGGCGAGTGGGGGCGCTCCCTCTTCCAAAAGCCCGCTTCCAGGTGAC 2173  
 OY 2084 ATCAGCTCTCTGAGGCTGACAGGCTCCCGGACGAGGAGCGGAGGTGAGTGC 2143  
 DB 2174 ATCAGCTCTCTGAGGCTGACAGGCTCCCGGACGAGGAGCGGAGGTGAGTGC 2233  
 OY 2144 ATCTACTCCGTCACCTTCACTCTCATCTGAGGTCCAGCGCTGAGTCAAGGAGTGTG 2203  
 DB 2234 ATCTACTCCGTCACCTTCACTCTCATCTGAGGTCCAGCGCTGAGTCAAGGAGTGTG 2293  
 OY 2204 GAGACATCTCAGGACACGCTCTGAGCACTCATGACAGCCCTCGTGAAGGCTGTGGA 2263  
 DB 2294 GAGACATCTCAGGACACGCTCTGAGCACTCATGACAGCCCTCGTGAAGGCTGTGGA 2353  
 OY 2264 GACGAGAAAGAGGGGGCCACAGACCGGCTGCTGTGCGCCACCCGAAAGCTGTGAGCC 2323  
 DB 2354 GACGAGAAAGAGGGGGCCACAGACCGGCTGCTGTGCGCCACCCGAAAGCTGTGAGCC 2413  
 OY 2324 CCACCGGCGCCCGACAGACCGAGCTGAGCACTTCCCGCGGAGGCGCCCGCAAGGAC 2383  
 DB 2414 CCACCGGCGCCCGACAGACCGAGCTGAGCACTTCCCGCGGAGGCGCCCGCAAGGAC 2473  
 OY 2384 AAGAAGCTCTGCGCACCAACGAGGAGCCCTGCGCCCTACACCGAGGGGCGGGGAGGA 2443  
 DB 2474 AAGAAGCTCTGCGCACCAACGAGGAGCCCTGCGCCCTACACCGAGGGGCGGGGAGGA 2533  
 OY 2444 GGGGACCCCTCCACCCCTTCCGTCGCGCCCAACTGTGAATCTGTAATTAAGGCCA 2503  
 DB 2534 GGGGACCCCTCCACCCCTTCCGTCGCGCCCAACTGTGAATCTGTAATTAAGGCCA 2593  
 OY 2504 AGGAACATGTGCGGAGGGGGGTGAGACAAAAACCGGCTTCCCTGACGAGGAGTGGCT 2563



QY	966	CGTGCCTGGAGACCGACCCCTGCGCCGGGATGAGCCATGCGGAGGCTGCATATCCAAAGGAG	1023
Db	781	CGTGCCTGGAGACCGACCCCTGCGCCGGGATGAGCCATATCGGAGGCTGCATATCCAAAGGAG	840
QY	1026	AGCTGGACCCCGAGCTCTTGAAGAGCATGCGATCACTGGGCTGCTTGACGAGACCGCGAGA	1085
Db	841	AGCTGGACCCCGAGCTCTTGAAGAGCATGCGATCACTGGGCTGCTTGACGAGACCGCGAGA	900
QY	1086	GGCTGCATCGCGAGCTCGGAGTGAAGAGAGAACCAAGAAAGATATATATATATATATCTGC	1145
Db	901	GGCTGCATCGCGAGCTCGGAGTGAAGAGAGAGAACCAAGAAAGATATATATATATATCTGC	960
QY	1146	TTTTGGATCGGAAGAGAGCGGATATCCACGTGTGAGAACACAGACCTCTCCGCCGGAATG	1205
Db	961	TTTTGGATCGGAAGAGAGCGGATATCCACGTGTGAGAACACAGACCTCTCCGCCGGAATG	1020
QY	1206	ATGTGTACCCCCCCCCGGAACCGTGTGGATTTCTCCATGCTGAGCGCTGCACGGGAAGCGCG	1265
Db	1021	ATGTGTACCCCCCCCCGGAACCGTGTGGATTTCTCCATGCTGAGCGCTGCACGGGAAGCGCG	1080
QY	1266	GACACAGAGCGGAAGTCCATGGAAGTCTCTGAGCATCATAGTCCGCGGGGGGTGGTGGCTCC	1325
Db	1081	GACACAGAGCGGAAGTCCATGGAAGTCTCTGAGCATCATAGTCCGCGGGGGGTGGTGGCTCC	1140
QY	1326	CTGTACCCACCCCAAGCGGCTTTGGAGATGGCCACACAGACCCAGAGATCCCGTAGCTCA	1385
Db	1141	CTGTACCCACCCCAAGCGGCTTTGGAGATGGCCACACAGACCCAGAGATCCCGTAGCTCA	1200
QY	1386	GTGGAGCCTTCACAGGGTCTGTCTCCAGAGCCCTTAAAGACACCCAAAGAGTCCGGTCTTTT	1445
Db	1201	GTGGAGCCTTCACAGGGTCTGTCTCCAGAGCCCTTAAAGACACCCAAAGAGTCCGGTCTTTT	1260
QY	1446	CCTTTTCACCGGAGCCGGGGGCTGTGAATGAGGCTCGAGCGGGGGCTCCCGCACTTCCA	1505
Db	1261	CCTTTTCACCGGAGCCGGGGGCTGTGAATGAGGCTCGAGCGGGGGCTCCCGCACTTCCA	1320
QY	1506	AAACGAGAGCGCTGACCTTTCGCGGGGGCCCAAGGGGTGGGGGGCGCCGGGAGACACCCCGCG	1565
Db	1321	AAACGAGAGCGCTGACCTTTCGCGGGGGCCCAAGGGGTGGGGGGCGCCGGGAGACACCCCGCG	1380
QY	1566	CCCCCAGTCCCGCTTCACACCCCTTGCAGCCCGCGCCGCCACAGGCTTCCCGCGCTCTGTGGG	1625
Db	1381	CCCCCAGTCCCGCTTCACACCCCTTGCAGCCCGCGCCGCCACAGGCTTCCCGCGCTCTGTGGG	1440
QY	1626	GGACCCCTTGGACGTGCGCTTGGACACAGCCCGCGGGCCAGTCCACACGGAGACCCCGGGGA	1685
Db	1441	GGACCCCTTGGACGTGCGCTTGGACACAGCCCGCGGGCCAGTCCACACGGAGACCCCGGGGA	1500
QY	1686	CAACACACACCCCGACGCGCGGCGGTGGGGGAGACCGCTGGAGAGTGGTGTCTGA	1745
Db	1501	CAACACACACCCCGACGCGCGGCGGTGGGGGAGACCGCTGGAGAGTGGTGTCTGA	1560
QY	1746	ATTCATATCCGCAACAGCTTCTCTGGGCTCCCTTCGCTTTCACCGGGCGCAAGATGACAGTCC	1805
Db	1561	ATTCATATCCGCAACAGCTTCTCTGGGCTCCCTTCGCTTTCACCGGGCGCAAGATGACAGTCC	1620
QY	1806	GTACCGCTGAGAGATGTCTACAGCTTGACGCGCAGAGTCCCGCGGAGCTGGGAAAAAGCT	1865
Db	1621	GTACCGCTGAGAGATGTCTACAGCTTGACGCGCAGAGTCCCGCGGAGCTGGGAAAAAGCT	1680
QY	1866	CGTGTTCGGGAACTTCACTTCTCTTGGACAAAGAAAGCAATATTTCTCTGCTCTAAAG	1925
Db	1681	CGTGTTCGGGAACTTCACTTCTCTTGGACAAAGAAAGCAATATTTCTCTGCTCTAAAG	1740
QY	1926	ACAAACCTTCAGAGCGATCCAAAGAGACATCTGTCCATATGCTTTTGTGTGATCCCAAGCC	1985
Db	1741	ACAAACCTTCAGAGCGATCCAAAGAGACATCTGTCCATATGCTTTTGTGTGATCCCAAGCC	1800
QY	1986	TGAGTCACAGTGTGTGTCTACAGACACACTTCAGGGCGAGTGAACAAGGCAATGTGCGGGC	2045
Db	1801	TGAGTCACAGTGTGTGTCTACAGACACACTTCAGGGCGAGTGAACAAGGCAATGTGCGGGC	1860

	Accession	Gene	Species	Length (nt)	GC Content (%)	Start Codon	Stop Codon	ORF Length (aa)	Protein Name	RefSeq ID	NCBI Gene ID	Ensembl Gene ID	Ensembl Protein ID	UniProt ID	Swiss-Prot ID	TrEMBL ID	Pfam ID	InterPro ID	GO Term	KEGG Pathway	KEGG Enzyme	KEGG EC	KEGG K01101	KEGG K01102	KEGG K01103	KEGG K01104	KEGG K01105	KEGG K01106	KEGG K01107	KEGG K01108	KEGG K01109	KEGG K01110	KEGG K01111	KEGG K01112	KEGG K01113	KEGG K01114	KEGG K01115	KEGG K01116	KEGG K01117	KEGG K01118	KEGG K01119	KEGG K01120	KEGG K01121	KEGG K01122	KEGG K01123	KEGG K01124	KEGG K01125	KEGG K01126	KEGG K01127	KEGG K01128	KEGG K01129	KEGG K01130	KEGG K01131	KEGG K01132	KEGG K01133	KEGG K01134	KEGG K01135	KEGG K01136	KEGG K01137	KEGG K01138	KEGG K01139	KEGG K01140	KEGG K01141	KEGG K01142	KEGG K01143	KEGG K01144	KEGG K01145	KEGG K01146	KEGG K01147	KEGG K01148	KEGG K01149	KEGG K01150	KEGG K01151	KEGG K01152	KEGG K01153	KEGG K01154	KEGG K01155	KEGG K01156	KEGG K01157	KEGG K01158	KEGG K01159	KEGG K01160	KEGG K01161	KEGG K01162	KEGG K01163	KEGG K01164	KEGG K01165	KEGG K01166	KEGG K01167	KEGG K01168	KEGG K01169	KEGG K01170	KEGG K01171	KEGG K01172	KEGG K01173	KEGG K01174	KEGG K01175	KEGG K01176	KEGG K01177	KEGG K01178	KEGG K01179	KEGG K01180	KEGG K01181	KEGG K01182	KEGG K01183	KEGG K01184	KEGG K01185	KEGG K01186	KEGG K01187	KEGG K01188	KEGG K01189	KEGG K01190	KEGG K01191	KEGG K01192	KEGG K01193	KEGG K01194	KEGG K01195	KEGG K01196	KEGG K01197	KEGG K01198	KEGG K01199	KEGG K01200	KEGG K01201	KEGG K01202	KEGG K01203	KEGG K01204	KEGG K01205	KEGG K01206	KEGG K01207	KEGG K01208	KEGG K01209	KEGG K01210	KEGG K01211	KEGG K01212	KEGG K01213	KEGG K01214	KEGG K01215	KEGG K01216	KEGG K01217	KEGG K01218	KEGG K01219	KEGG K01220	KEGG K01221	KEGG K01222	KEGG K01223	KEGG K01224	KEGG K01225	KEGG K01226	KEGG K01227	KEGG K01228	KEGG K01229	KEGG K01230	KEGG K01231	KEGG K01232	KEGG K01233	KEGG K01234	KEGG K01235	KEGG K01236	KEGG K01237	KEGG K01238	KEGG K01239	KEGG K01240	KEGG K01241	KEGG K01242	KEGG K01243	KEGG K01244	KEGG K01245	KEGG K01246	KEGG K01247	KEGG K01248	KEGG K01249	KEGG K01250	KEGG K01251	KEGG K01252	KEGG K01253	KEGG K01254	KEGG K01255	KEGG K01256	KEGG K01257	KEGG K01258	KEGG K01259	KEGG K01260	KEGG K01261	KEGG K01262	KEGG K01263	KEGG K01264	KEGG K01265	KEGG K01266	KEGG K01267	KEGG K01268	KEGG K01269	KEGG K01270	KEGG K01271	KEGG K01272	KEGG K01273	KEGG K01274	KEGG K01275	KEGG K01276	KEGG K01277	KEGG K01278	KEGG K01279	KEGG K01280	KEGG K01281	KEGG K01282	KEGG K01283	KEGG K01284	KEGG K01285	KEGG K01286	KEGG K01287	KEGG K01288	KEGG K01289	KEGG K01290	KEGG K01291	KEGG K01292	KEGG K01293	KEGG K01294	KEGG K01295	KEGG K01296	KEGG K01297	KEGG K01298	KEGG K01299	KEGG K01300	KEGG K01301	KEGG K01302	KEGG K01303	KEGG K01304	KEGG K01305	KEGG K01306	KEGG K01307	KEGG K01308	KEGG K01309	KEGG K01310	KEGG K01311	KEGG K01312	KEGG K01313	KEGG K01314	KEGG K01315	KEGG K01316	KEGG K01317	KEGG K01318	KEGG K01319	KEGG K01320	KEGG K01321	KEGG K01322	KEGG K01323	KEGG K01324	KEGG K01325	KEGG K01326	KEGG K01327	KEGG K01328	KEGG K01329	KEGG K01330	KEGG K01331	KEGG K01332	KEGG K01333	KEGG K01334	KEGG K01335	KEGG K01336	KEGG K01337	KEGG K01338	KEGG K01339	KEGG K01340	KEGG K01341	KEGG K01342	KEGG K01343	KEGG K01344	KEGG K01345	KEGG K01346	KEGG K01347	KEGG K01348	KEGG K01349	KEGG K01350	KEGG K01351	KEGG K01352	KEGG K01353	KEGG K01354	KEGG K01355	KEGG K01356	KEGG K01357	KEGG K01358	KEGG K01359	KEGG K01360	KEGG K01361	KEGG K01362	KEGG K01363	KEGG K01364	KEGG K01365	KEGG K01366	KEGG K01367	KEGG K01368	KEGG K01369	KEGG K01370	KEGG K01371	KEGG K01372	KEGG K01373	KEGG K01374	KEGG K01375	KEGG K01376	KEGG K01377	KEGG K01378	KEGG K01379	KEGG K01380	KEGG K01381	KEGG K01382	KEGG K01383	KEGG K01384	KEGG K01385	KEGG K01386	KEGG K01387	KEGG K01388	KEGG K01389	KEGG K01390	KEGG K01391	KEGG K01392	KEGG K01393	KEGG K01394	KEGG K01395	KEGG K01396	KEGG K01397	KEGG K01398	KEGG K01399	KEGG K01400
--	-----------	------	---------	-------------	----------------	-------------	------------	-----------------	--------------	-----------	--------------	-----------------	--------------------	------------	---------------	-----------	---------	-------------	---------	--------------	-------------	---------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------

**AUTHORS** Nagase, T., Nakayama, M., Nakajima, D., Kikuno, R. and Ohara, O.  
**TITLE** Prediction of the coding sequences of unidentified human genes. XX.  
 The complete sequences of 100 new cDNA clones from brain which code  
 for large proteins in vitro

**JOURNAL** DNA Res. 8 (2), 85-95 (2001)  
**MEDLINE** 2145130  
**REFERENCE** 2 (bases 1 to 2576)  
**AUTHORS** Ohara, O., Nagase, T. and Kikuno, R.

**TITLE** Direct Submission  
**JOURNAL** Submitted (27-Mar-2001) Osamu Ohara, Kazusa DNA Research Institute,  
 Department of Human Gene Research, 1532-3, Yena, Kisarazu, Chiba  
 292-0812, Japan (E-mail: cdna@kazusa.or.jp,  
 URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913,  
 Fax: 81-438-52-3914)

**FEATURES**  
**Source** Location/Qualifiers  
 1..2576

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="ha06731"  
 /tissue="brain"  
 /note="Vector: pBluescriptII SK plus"  
 1..2576  
 /gene="KIAA1811"  
 <..2149  
 /gene="KIAA1811"  
 /note="Start codon is not identified."  
 /codon\_start=2  
 /product="KIAA1811 protein"  
 /protein\_id="BAB47440.1"  
 /db\_xref="GI:14017839"  
 /translation="IVNREKLSSEVLMKVEREIAILKLEHPHYLKLDHYENKYL  
 LVEHVSQELFDYLVKRGRLPKRKARKEQIVSALDFCHSYISICHRDLKPEMLLD  
 EKNRIIADFGMASLQVDSLETSCSPHYACPEVKEKGRADMMSCGYILFA  
 LVLGALPFDDDDNROLLEKRYKGVFMHPHPIPDCCSLKGMIEVEEKRSLQIQIK  
 HWYLGKREKPPCLEPARGRYAMSLSPNGELDPVLESMASLGGCFRDERLHREL  
 KSEVNOEKMTIYLIDRKERYRSCEDODLPRNDNDPPKRYDPSMLSRGRKRRER  
 KSEVNLSTIDAGGGSPVPTKRALMAQHSORNSVSGASTGLSSFLSPRSRYEER  
 SEPEGADBARGGSPSTKQTLPPSPKRGVGADEPPPSARSTPLDPGSPRSRG  
 GTPHSLPHTPRASPTPGTPTPPSPGSGVGAAMRSRLNSINSLPSRFRKRM  
 QVPTAEEMSSLSPTSEPELAKRSMGNFISLDKEQIFLYLKDPLSIRKDIYHAF  
 SLPSLSHVSLSOTSEFAEKYASGSPVFOKRVPRVOVISSEGEPEPSRDRSGGGGI  
 YVTPTLISGSPSRFRVETIOAOLISTHDOPSYQALADEKNKNAQTRPNAGPRSLD  
 PPGRRDPELSSPSRPPKDKLLATNGPLP"

**BASE COUNT** 508 a 842 c 766 g 460 t  
**ORIGIN**

**Query Match** 86.4%; Score 2576; DB 9; Length 2576;  
**Best Local Similarity** 100.0%; Pred. No. 0;  
**Matches 2576; Conservative** 0; Mismatches 0; Indels 0; Gaps 0;

QY 274 GATCGTGACCGGAGAGCTGTGAGTGGTCTGTGATGAGGTGAGCGGGAGATCGC 333  
 DB 1 GATCGTGACCGGAGAGCTGTGAGTGGTCTGTGATGAGGTGAGCGGGAGATCGC 60  
 QY 334 CATCCGAACCTATGACCAACCAACATGTCCTCAAGCTCCACGAGCTACGAGAACAA 393  
 DB 61 CATCCGAACCTATGACCAACCAACATGTCCTCAAGCTCCACGAGCTACGAGAACAA 120  
 QY 394 GAATATTTTACCTGTTCTGAGCAAGTCTGGGGGGTNGATTTTCACTACTGCT 453  
 DB 121 GAATATTTTACCTGTTCTGAGCAAGTCTGGGGGGTNGATTTTCACTACTGCT 180  
 QY 454 AAGAGAGGAGACTGACGCCCAAGAGAGGCCGAAGATTCTTCCGCCAGATGTGCTGC 513  
 DB 181 AAGAGAGGAGACTGACGCCCAAGAGAGGCCGAAGATTCTTCCGCCAGATGTGCTGC 240  
 QY 514 GCTGAGCTTGGCCACAGCTCTCATCTGCACAGAGACCTAAGACCCGAGAACCTGCT 573  
 DB 241 GCTGAGCTTGGCCACAGCTCTCATCTGCACAGAGACCTAAGACCCGAGAACCTGCT 300  
 QY 574 TTTGGATGAGAAAAACAACATCCGATTGACAGACTTGGCATGGCGTCCCTGAGGTGG 633  
 DB 301 TTTGGATGAGAAAAACAACATCCGATTGACAGACTTGGCATGGCGTCCCTGAGGTGG 360

QY 634 GGACAGCTCTCTGAGAGCAGCTGCGGTCCGCCATTATGCTGTCCAGAGGTGATTA 693  
 DB 361 GGACAGCTCTCTGAGAGCAGCTGCGGTCCGCCATTATGCTGTCCAGAGGTGATTA 420  
 QY 694 GGGGAAAAATATGATGCGCCCGCGGAGACATGTGAGCTGTGAGTCTCTCTTCCG 753  
 DB 421 GGGGAAAAATATGATGCGCCCGCGGAGACATGTGAGCTGTGAGTCTCTCTTCCG 480  
 QY 754 CTTGCTGCTGGGGGGCTCTGCTCTTTGATGACGACAACTCCGCGAGTGTGAGAGT 813  
 DB 481 CTTGCTGCTGGGGGGCTCTGCTCTTTGATGACGACAACTCCGCGAGTGTGAGAGT 540  
 QY 814 GAAAGGGGCGCTTCCACATGCGCCCACTTCATCTCCAGATTGCCAGAGCTCTGAG 873  
 DB 541 GAAAGGGGCGCTTCCACATGCGCCCACTTCATCTCCAGATTGCCAGAGCTCTGAG 600  
 QY 874 GGAATGATCGAAGTGTGAGACCCCAAAAAAGGCTCAGCTGTGAGCAAAATTCAGAAATCC 933  
 DB 601 GGAATGATCGAAGTGTGAGACCCCAAAAAAGGCTCAGCTGTGAGCAAAATTCAGAAATCC 660  
 QY 934 TTGCTACTAGCGGGGAAACACAGACGACCCGCTGCTGAGCAGCCCTGGCCGCG 993  
 DB 661 TTGCTACTAGCGGGGAAACACAGACGACCCGCTGCTGAGCAGCCCTGGCCGCG 720  
 QY 994 GGTAGCCTAGCGGAGCCTGCTCCATCCAAAGAGAGCTGAGACCCGAGCTCTAGAGAGCAT 1053  
 DB 721 GGTAGCCTAGCGGAGCCTGCTCCATCCAAAGAGAGCTGAGACCCGAGCTCTAGAGAGCAT 780  
 QY 1054 GGCATACAGCGGCTGCTTCAGGGAGCGGAGAGGCTCATCGGAGCTGGCGAGTGAAGA 1113  
 DB 781 GGCATACAGCGGCTGCTTCAGGGAGCGGAGAGGCTCATCGGAGCTGGCGAGTGAAGA 840  
 QY 1114 GGAGAACCAAGAAAGATGATATATTATCTGCTTTTGGATCGAAGAGCGGTATCCAG 1173  
 DB 841 GGAGAACCAAGAAAGATGATATATTATCTGCTTTTGGATCGAAGAGCGGTATCCAG 900  
 QY 1174 CTGTGAGAGCAGAGAGCTGCTCCCGGAATGATGTTGACCCCCCGGAGAGCGTGTGGA 1233  
 DB 901 CTGTGAGAGCAGAGAGCTGCTCCCGGAATGATGTTGACCCCCCGGAGAGCGTGTGGA 960  
 QY 1234 TTTTCCCATGCTGAGCGCTGACGCGGAGGAGGAGCGGAGCGGAGGAGTCAATGAGTCT 1293  
 DB 961 TTTTCCCATGCTGAGCGCTGACGCGGAGGAGGAGCGGAGCGGAGGAGTCAATGAGTCT 1020  
 QY 1294 GAGCATACCGATGCGCGGGGTGTGCTCCCTGTACCCACCGAGGAGGCTTGTGAGAT 1353  
 DB 1021 GAGCATACCGATGCGCGGGGTGTGCTCCCTGTACCCACCGAGGAGGCTTGTGAGAT 1080  
 QY 1354 GGCCGAGCAGAGCAGAGATCCGCTAGCGTCAAGGAGGCTCCAGGGTCTGTCTCCAG 1413  
 DB 1081 GGCCGAGCAGAGCAGAGATCCGCTAGCGTCAAGGAGGCTCCAGGGTCTGTCTCCAG 1140  
 QY 1414 CCCTCTAAGCAGCCCAAGAGATCCGCTCTTCTTTCACCGGAGCGCGGGGCTGTGAGA 1473  
 DB 1141 CCCTCTAAGCAGCCCAAGAGATCCGCTCTTCTTTCACCGGAGCGCGGGGCTGTGAGA 1200  
 QY 1474 TGAAGCTGAGGCGGGGCTCCCGGACTTCCAAAACGAGACGCTGCTTCTCGGGGGCC 1533  
 DB 1201 TGAAGCTGAGGCGGGGCTCCCGGACTTCCAAAACGAGACGCTGCTTCTCGGGGGCC 1260  
 QY 1534 CAGGGGTGGGGGGCGCGGGGAGCAGCCCGCCCGGAGTCCGCGCTCCACACCCCTGCC 1593  
 DB 1261 CAGGGGTGGGGGGCGCGGGGAGCAGCCCGCCCGGAGTCCGCGCTCCACACCCCTGCC 1320  
 QY 1594 CGGCGCCCAAGGCTCCCGGCTCTCTTGTGGGGAAGCCCTTGCACCTGCTGACAC 1653  
 DB 1321 CGGCGCCCAAGGCTCCCGGCTCTCTTGTGGGGAAGCCCTTGCACCTGCTGACAC 1380  
 QY 1654 GCCCGGGGCGAGTCCACCGGAGACCCCGGGAGCAACACACCCCGGAGGCGGGGGTGG 1713  
 DB 1381 GCCCGGGGCGAGTCCACCGGAGACCCCGGGAGCAACACACCCCGGAGGCGGGGGTGG 1440

[illegible]

Db	2521	CCCCCAGACCTCTGTACGATTTGCTCTCCGGAAGAAATTCGTGTTCCGCTG	2576
RESULT 5			
LOCUS	AXI65526	2385 bp	DNA
DEFINITION	Sequence 17 from Patent WO0138503.		
ACCESSION	AXI65526		
VERSION	AXI65526.1	GI:14546871	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. (bases 1 to 2385)		
AUTHORS	Plomman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R., Plomman, P. and Clary, D.S.		
TITLE	Novel human protein kinases and protein kinase-like enzymes		
JOURNAL	Patent: WO 0138503-A 17 31-MAY-2001; Sugen, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..2385		
BASE COUNT	494 a 768 c 731 g 392 t		
ORIGIN			
Query Match	75.8%; Score 2259; DB 6; Length 2385;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 2259; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Y	164	CAGGCCCAATATGTGGGGCCCTATTCGGCTGGAAGAGCGTGGCAAGACAGAGGG	223
Db	127	CAGGCCCAATATGTGGGGCCCTATTCGGCTGGAAGAGCGTGGCAAGACAGAGGG	186
Y	224	CTGGTTAAACTCGGGGTCACATCGCATACGGGTGCAAGGTCGCCATCAAGATCGGAAC	283
Db	187	CTGGTTAAACTCGGGGTCACATCGCATACGGGTGCAAGGTCGCCATCAAGATCGGAAC	246
Y	284	CGGAGAGAGCTGTGAGTCGGAGTCGGTCGATGGAAGGTGAGACGGAGATCCCATCTGAA	343
Db	247	CGGAGAGAGCTGTGAGTCGGAGTCGGTCGATGGAAGGTGAGACGGAGATCCCATCTGAA	306
Y	344	CTCATGCAACACCCACATGTCCTCAAGTCCACAGACGCTACGAGAACAAAGAAATTTTG	403
Db	307	CTCATGCAACACCCACATGTCCTCAAGTCCACAGACGCTACGAGAACAAAGAAATTTTG	366
Y	404	TACCTGGTTCTGAGACAGCTCTCGGGGGGTAGCATTCGACTACCTGGTAAAGAGGGG	463
Db	367	TACCTGGTTCTGAGACAGCTCTCGGGGGGTAGCATTCGACTACCTGGTAAAGAGGGG	426
Y	464	AGAGCTACGCCCCCAAGAGGGCCGAAAGTTCTTCGCCACAGATTGTGTCTGCGTGAAC	523
Db	427	AGAGCTACGCCCCCAAGAGGGCCGAAAGTTCTTCGCCACAGATTGTGTCTGCGTGAAC	486
Y	524	TGCGACAGCTACTTCACATCTGCGACAGAGACCTTAAAGCCGGAACCTGTTTGGATGAG	583
Db	487	TGCGACAGCTACTTCACATCTGCGACAGAGACCTTAAAGCCGGAACCTGTTTGGATGAG	546
Y	584	AAAAACAACATCCGCAATTTGACAGACTTCGGCATGGCGCTCCCTGACAGTGGGGACAGC	643
Db	547	AAAAACAACATCCGCAATTTGACAGACTTCGGCATGGCGCTCCCTGACAGTGGGGACAGC	606
Y	644	CTGGAGACCAAGCTGGGGGTCCCCCATTTATGCGTGTCCAGAGGTGATTTAAGGGGAAAA	703
Db	607	CTGGAGACCAAGCTGGGGGTCCCCCATTTATGCGTGTCCAGAGGTGATTTAAGGGGAAAA	666
Y	704	TATGATGCGCGCGGGGACAGATGTGAGAGCTGTGAGAGTCATCTTGGCCCTGCTCGT	763
Db	667	TATGATGCGCGCGGGGACAGATGTGAGAGCTGTGAGAGTCATCTTGGCCCTGCTCGT	726
Y	764	GGGGCTCTGCGCTTTATGACGACCAACCTCCGACGCTGCTGAGAAAGTGAACGGGGC	823

Db 727 GGGGGCTCTCCCTTATGATGACGACAACTCCGCGAGCTGTGGAGAAAGGTAAACGGGGC 786  
 Oy 824 GTCTTCACATGCGCCCACTTCCTTCAGATTCAGAGCCCTCTGAGGGAGATGATC 883  
 Db 787 GTCTTCACATGCGCCCACTTCCTTCAGATTCAGAGCCCTCTGAGGGAGATGATC 846  
 Oy 884 GAATGAGAGCCCGGAAAGAGTCTGAGAGCAATTCAGAAACATCTTGTGACTTA 943  
 Db 847 GAATGAGAGCCCGGAAAGAGTCTGAGAGCAATTCAGAAACATCTTGTGACTTA 906  
 Oy 944 GGGGGGAAACAGAGGACGAGCCGCTGAGAGCAAGCCCTGCGCCCGGTGAGCCATG 1003  
 Db 907 GGGGGGAAACAGAGGACGAGCCGCTGAGAGCAAGCCCTGCGCCCGGTGAGCCATG 966  
 Oy 1004 CGAGGCTGCTCAACAGAGAGTGGAGCCCGGAGCTCTGAGAGCATGGATCACTG 1063  
 Db 967 CGAGGCTGCTCAACAGAGAGTGGAGCCCGGAGCTCTGAGAGCATGGATCACTG 1026  
 Oy 1064 GGGCTCTTCAAGGAGCGGAGAGGCTGATCGGAGCTGCGCACTGAGAGAGAGCAAA 1123  
 Db 1027 GGGCTCTTCAAGGAGCGGAGAGGCTGATCGGAGCTGCGCACTGAGAGAGAGCAAA 1086  
 Oy 1124 GAAAGATGATATATATATGCTTTTGGATCGGAAAGAGGGATCCAGCTGTGAGGAG 1183  
 Db 1087 GAAAGATGATATATATATGCTTTTGGATCGGAAAGAGGGATCCAGCTGTGAGGAG 1146  
 Oy 1184 CAGAGCTGCTCCCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1243  
 Db 1147 CAGAGCTGCTCCCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1206  
 Oy 1244 CTGAGCGCTGACGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 1303  
 Db 1207 CTGAGCGCTGACGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 1266  
 Oy 1304 GATGCGGAGGAGTGTGGCTCCCTGTCACACCGGAGGAGGCTTGGAGATGGCCAGGAG 1363  
 Db 1267 GATGCGGAGGAGTGTGGCTCCCTGTCACACCGGAGGAGGCTTGGAGATGGCCAGGAG 1326  
 Oy 1364 AGCAGAGATCCGCTAGCGTCACTGAGAGCTCAACGAGGCTCTGCTCAGCCCTTAAGC 1423  
 Db 1327 AGCAGAGATCCGCTAGCGTCACTGAGAGCTCAACGAGGCTCTGCTCAGCCCTTAAGC 1386  
 Oy 1424 AGCCCAAGAGTCCGCTCTTCTTCTTCAACCGGAGCGGAGCGGAGCGGAGCGGAGCG 1483  
 Db 1387 AGCCCAAGAGTCCGCTCTTCTTCTTCTTCAACCGGAGCGGAGCGGAGCGGAGCG 1446  
 Oy 1484 GGGGGGAGTCCCGGAGTCCCAAAACGAGAGCTGCTTCTGAGGGGCGCCAGGGGTGG 1543  
 Db 1447 GGGGGGAGTCCCGGAGTCCCAAAACGAGAGCTGCTTCTGAGGGGCGCCAGGGGTGG 1506  
 Oy 1544 GGGGGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAG 1603  
 Db 1507 GGGGGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAGTCCCGGAG 1566  
 Oy 1604 GGGTCCCCGGGCTCTTGTGGGAGGAGCCCGCTGAGTCTGCTTGCACACGCGCGGG 1663  
 Db 1567 GGGTCCCCGGGCTCTTGTGGGAGGAGCCCGCTGAGTCTGCTTGCACACGCGCGGG 1626  
 Oy 1664 AGTCCCAACGAGAGCGGAGGAGCAACACGAGCCCGGAGCGGAGGAGGAGGAGGAG 1723  
 Db 1627 AGTCCCAACGAGAGCGGAGGAGCAACACGAGCCCGGAGCGGAGGAGGAGGAGGAG 1686  
 Oy 1724 GCGGCTGAGAGTCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1783  
 Db 1687 GCGGCTGAGAGTCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1746  
 Oy 1784 CACCGGAGAGTCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1843  
 Db 1747 CACCGGAGAGTCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1806  
 Oy 1844 TCCCGGAGAGTCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1903

Db 1807 TCCCGGAGTCTGAGAGAGTCTCTGAGAGTCTCTGAGAGTCTCTGAGAGAGAG 1866  
 Oy 1904 CAATATATCTGCTGTGTAAAGAGCAAACTCTCAGAGATCAAGAGAGATGCTCAT 1963  
 Db 1867 CAATATATCTGCTGTGTAAAGAGCAAACTCTCAGAGATCAAGAGAGATGCTCAT 1926  
 Oy 1964 GCTTTCTGT 2023  
 Db 1927 GCTTTCTGT 1986  
 Oy 2024 GAGTACAGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2083  
 Db 1987 GAGTACAGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2046  
 Oy 2084 ATCAGCTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2143  
 Db 2047 ATCAGCTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2106  
 Oy 2144 ATCTACTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2203  
 Db 2107 ATCTACTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2166  
 Oy 2204 GAGACATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2263  
 Db 2167 GAGACATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2226  
 Oy 2264 GAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2323  
 Db 2227 GAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2286  
 Oy 2324 CCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2383  
 Db 2287 CCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2346  
 Oy 2384 AAGAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2422  
 Db 2347 AAGAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2385

RESULT 6  
 BC016681  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens.  
 Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 2128)  
 Strausberg, R.  
 Direct Submision  
 Submitted (31-OCT-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Lou Staudt  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [villalonbcm.tmc.edu](mailto:villalonbcm.tmc.edu)  
 Villalon, D.K., Luna, R.A., Hale, S.M., Huliy, S., Lu, X., Garcia,  
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
 Muzny, D.M., Gibbs, R.A.

REMARK  
 COMMENT  
 Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRK Plate: 26 Row: 3 Column: 5  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis.

## FEATURES

## source

## Location/Qualifiers

1..2128  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:21870 IMAGE:434183"  
 /tissue\_type="Lymph, Lymphoma"  
 /clone\_1ib="Nih\_MGC\_85"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6"  
 227..1648  
 /codon\_start=1  
 /product="Unknown (protein for MGC:21870)"  
 /protein\_id="AAH16681.1"  
 /db\_xref="GI:16741776"  
 /translation="MRSLEPSELDPELVESNASLIGCFDREHRLHRESENOEKM  
 IYLLDKRERYPCEDDLPKNDVDPKRYDPSMLSHGRERPERKMEYLSD  
 AGGSGPVPTRALMAHOHORSRVSASGTSSPSLSPSPSPSPSPSPSPSP  
 RGGSPSTKTOTLPKSPRGGAGGAPPPSPSPSPSPSPSPSPSPSPSPSP  
 PRAPSTGTCTPPPPSPGGGAGGAMRSLNRSFLSPSPSPSPSPSPSPSP  
 LTPESPRLAKSMFNGNTSLDKREDFIVLKDPLSLKADIVHAFSLPSLSHVL  
 SOTSRRAETKASGGSPVRFQVDFVDFVDFVDFVDFVDFVDFVDFVDFV  
 PSRRKRVETIOAQLDLSHDHPDVSVALDERKGAQTRPAGAPRPSLPDPDP  
 SSSPRAPPKDKLLATNCTPLP"

## CDS

BASE COUNT 408 a 707 c 651 g 362 t  
 ORIGIN

Query Match 65.6%; Score 1953.6; DB 9; Length 2128;  
 Best Local Similarity 99.8%; Pred. No. 0;

Matches 1956; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

943 AGGCGGGAACACGAGCCGAGCCGCTGCTGAGACGACCCCTGCGCGCGCGGATGACAT 1002  
 169 AGGCGGGAACACGAGCCGAGCCGCTGCTGAGACGACCCCTGCGCGCGCGGATGACAT 228  
 1003 GCGGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062  
 229 GCGGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 288  
 1063 GCGGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122  
 289 GCGGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 348  
 1123 AGAAAAAGATGAT 1182  
 349 AGAAAAAGATGAT 408  
 1183 CCAGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1242  
 409 CCAGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 468  
 1243 GCTGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1302  
 469 GCTGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528  
 1303 CGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1362  
 529 CGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 588  
 1363 CAGCCAGAGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1422  
 589 CAGCCAGAGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648  
 1423 CAGCCAGAGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1482  
 649 CAGCCAGAGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708  
 1483 AGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1542

709 AGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 768  
 1543 GGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1602  
 769 GGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 828  
 1603 AGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1662  
 829 AGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 888  
 1663 CAGTCCACCGGAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1722  
 889 CAGTCCACCGGAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 948  
 1723 AGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1782  
 949 AGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1008  
 1783 TCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1842  
 1009 TCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1068  
 1843 CTCGCCGAGAGCTGGCAAAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1902  
 1069 CTCGCCGAGAGCTGGCAAAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1128  
 1903 ACAATATATCTCGTGTAAAGAGCAAACTCTGACAGATCAAAAGCAATCTGCTGCA 1962  
 1129 ACAATATATCTCGTGTAAAGAGCAAACTCTGACAGATCAAAAGCAATCTGCTGCA 1188  
 1963 TGCCCTTCTGCTGATCCCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2022  
 1189 TGCCCTTCTGCTGATCCCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1248  
 2023 CGAGTACAAAGGCGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2082  
 1249 CGAGTACAAAGGCGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1308  
 2083 CATCAGCTCTCTGATGAGGCTCCAGAGCCCTCCCGGCGGAGCGGAGCGGAGGAGTGTG 2142  
 1309 CATCAGCTCTCTGATGAGGCTCCAGAGCCCTCCCGGCGGAGCGGAGCGGAGGAGTGTG 1368  
 2143 CATCAGCTCTCTGATGAGGCTCCAGAGCCCTCCCGGCGGAGCGGAGCGGAGGAGTGTG 2202  
 1369 CATCAGCTCTCTGATGAGGCTCCAGAGCCCTCCCGGCGGAGCGGAGCGGAGGAGTGTG 1428  
 2203 GGAAGCATTCAGAGGAGCTGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAG 2262  
 1429 GGAAGCATTCAGAGGAGCTGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAG 1488  
 2263 AGAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2322  
 1489 AGAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1548  
 2323 CCCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2382  
 1549 CCCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1608  
 2383 CAAAGAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2442  
 1609 CAAAGAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1668  
 2443 AGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2502  
 1669 AGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1728  
 2503 AAGGAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2562  
 1729 AAGGAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1788  
 2563 TCACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2622  
 1789 TCACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1848

QY	1882	CATCTCCTTGGACAAAGAGAACAATATTCCTCGTGCTCTAAAGACAACCTCTCAGCAG	1941
Db	1	CATCTCCTTGGACAAAGAGAACAATATTCCTCGTGCTCTAAAGACAACCTCTCAGCAG	60
QY	1942	CATCAAAACACACATATCGACATAGCCTTCCTGCGATGCCCAAGCCGAGTGCACAGTGTCT	2001
Db	61	CATCAAAACACACATATCGATGCTTCTGTGCGATGCCCAAGCCGAGTGCACAGTGTCT	120
QY	2002	GTCAACAGACAGCTTTCAGGGGCCGAGTACAAAGGCCAGTGGGGCCCTCGTCTCCAAAA	2061
Db	121	GTCAACAGACAGCTTTCAGGGGCCGAGTACAAAGGCCAGTGGGGCCCTCGTCTCCAAAA	180
QY	2062	GCCCGTCCGCTTCCAGGTGTGAGACATACGCTCCCTGTGAGGGTCCAGAGCCCTCCCGCGAGC	2121
Db	181	GCCCGTCCGCTTCCAGGTGTGAGACATACGCTCCCTGTGAGGGTCCAGAGCCCTCCCGCGAGC	240
QY	2122	GGACGGCAGCGAGGTGGTGGCATCTACTCCGTCACTTCATCTCATCTCGGGTCCAG	2181
Db	241	GGACGGCAGCGAGGTGGTGGCATCTACTCCGTCACTTCATCTCATCTCGGGTCCAG	300
QY	2182	CCGTGCGTTCAAGCGATGTGTGAGACCATCCAGGCACAGCTCTGTGACATCTCATGACCA	2241
Db	301	CCGTGCGTTCAAGCGATGTGTGAGACCATCCAGGCACAGCTCTGTGAGACATCATGACCA	360
QY	2242	GCCCTCCGTCGACGGCCCTGGGCGAGAGAGAAAGGGGGCCCGAGCCGGCCCTGCTGTGC	2301
Db	361	GCCCTCCGTCGACGGCCCTGGGCGAGAGAGAAAGGGGGCCCGAGCCGGCCCTGCTGTGC	420
QY	2302	CCCACCCCGAAACCTCTGACAGCCCCCAACCCGGCCGCCAGACCCAGAGCTGAGACGCTTCC	2361
Db	421	CCCACCCCGAAACCTCTGACAGCCCCCAACCCGGCCGCCAGACCCAGAGCTGAGACGCTTCC	480
QY	2362	CCGCGGAGGGCCCCCCCAAGAGACAAGAAGTCTGTGGGCACCAAGGAGGCCCTGTGGCCCTG	2421
Db	481	CCGCGGAGGGCCCCCCCAAGAGACAAGAAGTCTGTGGGCACCAAGGAGGCCCTGTGGCCCTG	540
QY	2422	ACCCACAGGGGGCGGGGAGGAGAGAGGGGAGCCCCCTCCACCCCTTCCTCGTGGCCCCCAACT	2481
Db	541	ACCCACAGGGGGCGGGGAGGAGAGAGGGGAGCCCCCTCCACCCCTTCCTCGTGGCCCCCAACT	600
QY	2482	GTGAATCTGTAAATTAAGGCCCAAGGACAATGTGCGGAGGGGGGTGAGACAAAAAACCGGC	2541
Db	601	GTGAATCTGTAAATTAAGGCCCAAGGACAATGTGCGGAGGGGGGTGAGACAAAAAACCGGC	660
QY	2542	CTTGGCCGTGACAGGAGTGGGGGTCCACAGGCGCTGCCCACTGGGGGTGGTTCTAGGGGAA	2601
Db	661	CTTGGCCGTGACAGGAGTGGGGGTCCACAGGCGCTGCCCACTGGGGGTGGTTCTAGGGGAA	720
QY	2602	CAGGGGGCGGGGAGAGCTGTTCTAATTTAATTAATGATTAATTAATTAATTAATTAATTTG	2661
Db	721	CAGGGGGCGGGGAGAGCTGTTCTAATTTAATTAATGATTAATTAATTAATTAATTAATTTG	780
QY	2662	ATCAATCTCTTCGCGGGGTGGGGTGGGGGACAGGAGCGGAGACTGTGTGGGGTGGCTTAGCA	2721
Db	781	ATCAATCTCTTCGCGGGGTGGGGTGGGGGACAGGAGCGGAGACTGTGTGGGGTGGCTTAGCA	840
QY	2722	GATCGGAGACAGGACCTCTGTCCCTGTGTGTCGCCCAACCCCTCTCTCCGGGCCCTCC	2781
Db	841	GATCGGAGACAGGACCTCTGTCCCTGTGTGTCGCCCAACCCCTCTCTCCGGGCCCTCC	900
QY	2782	TCCCTGTGTCTTCCGCCCAAGACTTCTGTACGGAATTGTCTCCGGAAGAAATTTCGT	2841
Db	901	TCCCTGTGTCTTCCGCCCAAGACTTCTGTACGGAATTGTCTCCGGAAGAAATTTCGT	960
QY	2842	TTTCGGCTGA 2850	



Query Match	32.2%	Score 961;	DB 6;	Length 2219;
Best Local Similarity	70.1%	Pred. No. 4,7e-171;		
Matches 1477;	Conservative	0;	Mismatches 435;	Indels 195;
				Gaps 6

QY	159	CCGACGAGCCCAATATGTGGGCCCTTATCGGCTGGAGAAGACGCTGGGCAAAAGGACCA	218
Db	81	CGCACACGCGCAGTATGTTGGGCCCTTACCGGCTGGAGAAGACGCTGGGCAAGGGGGAGA	140
QY	219	CAGGCGTGTTAACTCGGGGTCCTACTGCATCACGGGTGAGAAAGTGCCTATCAAGATCG	278
Db	141	CAGGCTGTGTGAAGCTGGGGGGTTTCACTGGGTCCCTGACCAGAAAGGTGGCCATCAAGATCG	200
QY	279	TGAACCGGAGAAAGCTGTGGAGTGCSTGCTGATGTAAGGTGGAGCGGGAGATCGCCATCC	338
Db	201	TCAACCGTGAAGGCTCACCGAGTCCGCTGTGATGAAGGTGGAGCGGGAGATCGCCATCC	260
QY	339	TGAAGCTCATGACACCCACATATGTCCTCAAGCTCCACGACGCTCTACGAGAACAAAGAT	398
Db	261	TGAAGCTCATGAGCACCCCCACGTCCTTAAAGCTGCACGACGCTTATGAACAAACAAAAT	320
QY	399	ATTGTACTGTGTTCTGGAGACACGTCCTCGGGGGGTAGAGTAAATTCGACTACTGGTAAAGA	458
Db	321	ATTGTACTGTGTTGCTAGAACACGTCGTACAGTGTGAGTGTAGCTCTTGACTACTGGTGAAGA	380
QY	459	AGGGAGACGTGACGCCCAAGSAGAGGCCGAAAGTCTTCCCCACAGATTGTGCTCGCGTGG	518
Db	381	AGGGAGGCTGAGCGCTAAGSAGAGCTGAGAAAGTCTTCCGGGAGATCATCTCTCGCGTGG	440
QY	519	ACTTCTGCCACAGCTACTCCTCATCTGCCACAGAGACCTAAAGCCCGAGACCTGCTTTTGG	578
Db	441	ACTTCTGCCACAGCCACTCCTCATATGCTACAGAGGATGAAACCTGAAACCTCTCGCTGG	500
QY	579	ATGAGAAAAACAACATCCGATTTGACAGACTTGGGCAATGGGTCCTCAGGTGGGGGACA	638
Db	501	ACGAGAGAACACATCCGATTCGACAGACTTTGGCATGGGTCCTCAGGTGGGGGACA	560
QY	639	GCCTCTGAGAACACAGCTGGGGTCCGCCATTATATGCGTGTCCAGAGTGAATTAAGGGG	698
Db	561	GCCTGTGGAGAACACACTGTGGGTCCGCCATTACAGCTTCCCGGAGGTGATCCGGGGGG	620
QY	699	AAAAATATATGGCCCGCGGCGAGACATGGAGCTGTGAAGTCAATCCTCTGGCCCTGGC	758
Db	621	AGAAATATGACCGCGCGGAGAGGCGAGCTGGAGACTCGGCGTATCTCTGTCGCTTGGC	680
QY	759	TGCTGGGGGCTCTGCCCTTTGATGACGACACAACCTCCGCCAGCTGCTGGAGAAGGTGAAC	818
Db	681	TGCTGGGGGCTCTGCCCTTTCGACGATACACAACCTTGACAGACGCTGCTGGAGAAGGTGAAC	740
QY	819	GGGGGCTCTTCCACATGGCCCACTTCAATTTCTCCAGATTTGCCAAGACCTCTCTGAGGGAA	878
Db	741	GGGGGCTCTTCCACATGGCCCACTTATATCCCGCCGACTGCGACAAGTCTGTACGGGGCA	800
QY	879	TGATCAAGTGAAGGCCCGAAGAAAAAGGCTCAAGTCTGAGAGCAAAATTCAGAAAAATCCTTGGT	938
Db	801	TGAGCGCAGGTGGACGCGCGACGCGCCTCAACGCTAGAGACACATTTCAAGAACACATATGGT	860
QY	939	ACCTAGCGGGGAAACAAGAGCCAGACCCGTGCTGGAGCCAGCCCTGGCCCGCGGGTAG	998
Db	861	ATATAGGGGGCAAGATATGAGCCCGAAGC-----AGAGCAGCCCATTTCTCTGCAAGGTGC	914
QY	999	CCATGGGAGACCTGCATCCAAAGSAGAGGTGAGACCCCGGAGCTGCTGAGAGCAATGGAT	1058
Db	915	AGATTCGCTGCTGGCCAGCTGTGAAGACATTCGACCCCGAGCTGTGACACAGCATGACT	974
QY	1059	CAGTGGGCTGCTTCAGGAGACCGGAGAGGCTGCATCGCAGCTCGCAGTGAAGAGGAGAGA	1118
Db	975	CAGTGGGCTGCTTCCGAGACCGCAACAAGTGTGCGAGGACCTGTGTCGAGGAGGAGA	1034
QY	1119	ACCAAGAAAAAGATGATATATATCTGCTTTTGGATTCGGAAGGAGACGATATCCCAAGCTGTG	1178
Db	1035	ACCAAGAGAGAGATGATTTACTTCTCTCTGTGACGCGAAGAAAGAAAGTATCCCGAGCCAGG	1094
QY	1179	AGGACAGAGACCTGCTTCCCGGAATATGTTGATCCCGCCCGGAGAGGCTGTGATTTCTC	1238
Db	1095	AGGATATAGAGACTGTCCCGCCCGGAGAGAGATATAGACCTCCCGCGGAAGGCTGTGACTGCC	1155
QY	1239	CCATGCTGAGCCGCTCACGGGAGAGCGCGACCGAGCGGAGAGTCCATGGAAGTCTGTGACA	1296

```

Db 1155 CGATGCTGAACCGGCAAGGAGGCGCCAGAACCAATTCATGAGTGTCTCACCG 1214
QY 1299 TCACGATGCGGGGGTGTGTGCTCCCTGTACCCACCCGAGGCGCTTGGAGTGGCC 1358
Db 1215 TGAC-----GAGGGGGGCTCCCGGCTGCTGGCGGGGCGCATAGATGGCC 1265
QY 1359 AGCAGACCCAGAGATCCGTAGCGTCAGTGAAGCTCCAGGGTGTCTCCAGCCCTC 1418
Db 1266 ACCACGGCCAGAGTGTCTGCTCATCAGCGGTGCTCTCAGGCTTCCACAGCCAC 1325
QY 1419 TAAGCAGCCCAAGAGTCCGGTCTTTTCCTTTCACCGGAGCCGGGGCTGAGATGAG 1478
Db 1326 TAGACGACCCCGC----- 1338
QY 1479 CTCGAGGCGGGGGCTCCCGACTTCCAAAACGAGACGCTGCTTCTGCGGGCCCAAGG 1538
Db 1339 ----- 1338
QY 1539 GTGGGGGGGGGGGAGAGACCCCGCCCGCAGTCCCGCTCCACACCCCTGCCCGCC 1598
Db 1339 -----GGTACCCCTCACCCCTCACCAAGGGGAGTCCCTCCCG----- 1378
QY 1599 CCCAAGGCTCCCGGCTCTCTGGGGGAGACCCCGCTTGAAGTGGCTGACACGCGCC 1658
Db 1379 -----ACCCCAAGGGGAGACCTGTCTCCACAGCCAA 1409
QY 1659 GGGCAGTCCCAACCGGGACCCCGGGGACAAACACACCCCGAGCCCGGGGCTG 1718
Db 1410 AGGAGAGCCCGGCTGGCAGCCCAACCCCGCGCGCTCCAGCCCG-----AGCGTGC 1463
QY 1719 GGGGAGCGCGCGGAGAGTCTCTCACTCCAGTCCGCAACAGCTTCTGCGCTCCCTC 1778
Db 1464 GAGGGGTGCGCTGGAGGGCGCGGCTCACTCCAGTCCAAAGACCTTCTGCGCTCACCC 1523
QY 1779 GCTTTCACCGGCGCAAGATGCTCCCTACCGCTGAGAGATGTCCAGCTTGCAGCGCAG 1838
Db 1524 GCTTTCACCGGCGGAAACGCAAGTTCGACGCGCGGAGAGATGTCCAACTTGCAGCAG 1583
QY 1839 AGTCTCCCGGAGTGGCAAAAGCTCTGTGTTGGGAACTTCAATCTCTTGGAGCAAG 1898
Db 1584 AGTCTCCCGGAGTGGCAAAAGCTCTGTGTTGGGAACTTCAATCTCTTGGAGCAAG 1643
QY 1899 AAGCAAAATATTCCTGCTGCTAAAGAGCAAACTCTCAGCAGATCAAGAGCAATCG 1958
Db 1644 AGGAGCAGATCTTGTGCTATCAAAAGCAAACTCTGAGCTCCATCAAGGCTGACATCG 1703
QY 1959 TCCATGCTCTGTGATGCTCCAGCTGAGTCAAGTGTGTGTACAGACCACTTCA 2018
Db 1704 TGCAGGCTCTGTGATGCTCCAGGCTCAGGCTCAGCTCCCAAAACGAGCTTCC 1763
QY 2019 GGGCGGAGTACAGGCGAGTGGCGGCGCTCGTCTTCCAAAAGCCCGCTCCAGG 2078
Db 1764 GGGCGGAGTACAGGCGAGTGGCGGCGCTCGTCTTCCAAAAGCCCGCTCCAGG 1823
QY 2079 TGCAGATGAGTCTCTGTGCTGAGGCTCCAGGCGCTCCCGGAGGAGCGAGGAGT 2138
Db 1824 TTGAATATCACTACAGAGAGGT-----GGGAGGGGCGCAGAGAGAGA 1865
QY 2139 GTGCGATCTACTCTGCTACCTTCACTCTCATCTCGGCTCCAGCGCTCGGTTCAAGGAG 2198
Db 1866 ACGGATCTACTCTGCTACCTTCACTCTCATCTCGGCTCCAGCGCTCGGTTCAAGGAG 1925
QY 2199 TGTGTGAACCATTCAGGACAGCTCTGAGCACTATCAAGCCCTCGTGCAGGCGCC 2258
Db 1926 TGTGTGAACCATTCAGGACAGCTCTGAGCACTATCAAGCCCTCGTGCAGGCGCC 1985
QY 2259 TGGGAGA 2265
Db 1986 TGTGAGA 1992

```

```

HSA6701
LOCUS HSA6701 1956 bp mRNA linear PRI 14-MAR-2001
DEFINITION Homo sapiens mRNA for putative serine/threonine protein kinase.
ACCESSION AU006701
VERSION AU006701.1 GI:3217027
KEYWORDS putative; serine/threonine protein kinase.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1956)
AUTHORS Stanchi,F., Bertocco,E., Toppo,S., Dioguardi,R., Simionati,B.,
Cannata,N., Zambello,R., Lanfranchi,G. and Valle,G.
TITLE Characterization of 16 novel human genes showing high similarity to
Yeast sequences
JOURNALS Yeast 18 (1), 69-80 (2001)
MEDLINE 21064499
PUBMED 11124703
REFERENCE 2 (bases 1 to 1956)
AUTHORS Stanchi,F.
TITLE Direct Submision
JOURNALS Submitted (02-JUN-1998) Stanchi F., CRIBI Biotechnology Centre,
Universita' di Padova, Via G. Colombo 3, 35121 Padova, 35121, ITALY
FEATURES
source location/Qualifiers
1..1956
/organism="Homo sapiens"
/db_xref="taxon:9606"
/sex="female"
/tissue_type="whole brain"
/clonetype="clonetype: infant brain INTB"
/devstage="73 days post natal"
<1..1814
/codon_start=3
/product="putative serine/threonine protein kinase"
/protein_id="CA07196.1"
/db_xref="GI:3217028"
/db_xref="SPTREMBL:O60843"
/translacion="LIEHPHVLKLDHVENKRYLYLVEHVSSELPYLYKGRILP
KEARFEOILSALDFCHSHICHDLKPEMLIDENKIRIRADGMSILOYDSILE
TSGSPHYACEVIRGEKDYDRKADYSCVITLALYGLALPFDNDNLROLLEKVRG
VEHNPRIPTPCQSLRMSSEVDARRRLLELHONHWTIGSKNPEPEPPIPRVVOI
RSLSDLEDIDVDLSDMSLDFRNRLDLDLSEBNOEKMYFLDLDRKRPDQ
EDBLPPRNEIDLPKRVDSPLNLRGRKRPKMEVLSYDGGSPAPARATMAQ
HGORSISISGASGLSTPLSPRYTPHSPRSGSLPPKGPVHTPRESPAGTNPPT
PSSPSYGVPMARLINSIKNSFLSPFHRKLOVPPPEESNLTPSSPELAKSW
EGNFSLEKEIOIFVYIKDPLSLKAIIVAFSLSPLSHSVTSOTSRAXKXKRG
PAVROKPVKRVODITVYMGGAOKENGIVSTFTLLSGPSRFRKRVETIOQLLSTH
DPPAHOHSEPPPAPEGLSMGAGLKGQAVATSYESSL"
BASE COUNT 427 a 640 c 579 g 310 t
ORIGIN
Query Match 27.4%; Score 818; DB 9; Length 1956;
Best Local Similarity 68.6%; Pred. No. 4.5e-144;
Matches 1319; Conservative 0; Mismatches 410; Indels 195; Gaps 6;
QY 342 AGCTCATGGAACCCACATGCTCTCAAGCTCCAGAGCTCTACGAAACAATATT 401
Db 1 AGCTCATGGAACCCACACGCTCTAAAGCTCAGAGCTTATGAAACAAATAATT 60
QY 402 TGTACCTGCTGTGAGAGACGCTCTGGGGGTGAGCTATTGCACTACCTGTPAAGAG 461
Db 61 TGTACCTGCTGTGAGAGACGCTCTGGGGGTGAGCTATTGCACTACCTGTPAAGAG 120
QY 462 GAGAGCTAGCCGCAAGAGAGGCGGAAAGTTCTCCGCGAGATGTGTGGCTGCACT 521
Db 121 GAGAGCTAGCCGCTAAGAGAGGCTGGAAGTTCTCCGCGAGATCACTCTGCTGCACT 180
QY 522 TCTGCAAGCTACTCTGCTGCAAGAGAGCTTAAAGCCGAGAGAGCTTTTGATG 581
Db 181 TCTGCAAGCTACTCTGCTGCAAGAGAGCTTAAAGCCGCTGCTGAGAG 240
QY 582 AGAAAAACACATCGCACTTGGGATGCGCTCCCTGAGAGTGGGGAGAGCC 641

```

Db	241	AGAAAGAACATCCCGCATTCGAGACTTTGGCATGGCCGTCCCTCGACAGTTGGGACAGCC	300
Oy	642	TCCTGGAGACCAAGCTGGCGGTCGCCCATTAATGCGTGTCCAGAGTGAATTAAGGGGAAA	701
Db	301	TGTTGGAAACCAAGCTGTGGGTCCCCCACCACCTACGCTGCCCGGAGTGATCCGGGGAGGA	360
Oy	702	AATATGATGGCGCGCGGGCAGACATGTGGAGCGTGTGGAGTGCATCCCTTCGGCCCTGTGCG	761
Db	361	AGTATGACGGCGGGAAGGCGGAGCGAGCTGGAGCTGCGGCTCATCTGTTGCTTGTCTGG	420
Oy	762	TGGGGGCTCTGCGCTTTGTATGACGACAACTCCGCCAGCTGCTGAGAAAGTGAACGGG	821
Db	421	TGGGGGCTCTGCGCTTTCGACGATGACAACTTGCGACAGCTGCTGAGAAAGTGAACGGG	480
Oy	822	GGGTCTTCACATGCCCCACCTTCATTCCTCCAGATTGCCCAGAGCCCTCTGAGGGGAATGA	881
Db	481	GGGTCTTCACATGCGCGCATTTATCCCGCCGACGCGCAAGATGCTATCGGGGGCTGA	540
Oy	882	TCGAATGGAGCGCCCAAAAAAGGCTCAGTCTGGAGCAAAATTCAGAAACATCTTGTATAC	941
Db	541	GGAGATGTGACCGCCGACCGCGCGCTCAGCTATGAGACATTCAGAAACATATMTGGTATA	600
Oy	942	TAGCGCGGAACACAGACCCAGACCCGCTGCTGGAGCCAGCCCTCGGCCCGCGGGTACCA	1001
Db	601	TAGGGGGCAAGAAATAGCCCGCAAC-----AGAGAGGCCCATCTTCCTCGCAAGGTGAGA	654
Oy	1002	TGCGGAGCGCTGCATATCCAAACGGAGAGACTGGACCCCGCAGCTCTAGAGAGATGATGATCAC	1061
Db	655	TCCGCTCGCTGCGCCACGCTTGAGAGACATGACCCCCGACGCTGTGGACAGCATGCATCAC	714
Oy	1062	TGGGCTGCTTCAGGAGCCGCGAGAGGCTGCATGCGAGCTGCGCAGTGGAGAGAGAAC	1121
Db	715	TGGGCTGCTTCAGGAGCCGCAACAAAGCTCTGAGAGACCTGCTGTCCGAGAGAGAAAC	774
Oy	1122	AAGAAAGATGATATATTTATCTTGTCTTTGGATGGGAAGGCGGTATMCCAGCTGTGATG	1181
Db	775	AGGAGAGATGATTTACTTCTCTCTCTGTGACCCGGAAGAAAGATACCCAGCCAGAGAG	834
Oy	1182	ACCAAGACCTGCTCTCCCGGAAATGATGTTGACCCCCCGCGGAAGCGTGTGATTTCTCCA	1241
Db	835	ATGAGAGACCTGCTCTCTCTCTCTCTGTGACCCGGAAGAAAGATACCCAGCCAGAGAG	894
Oy	1242	TGCTAGCGCTACAGGGGAAGCGGCGACACAGAGCGGAAGTCCATATGAAAGTCTGTAGCATCA	1301
Db	895	TGCTGAACCGGACGCGCAAGCGGCGGCGCAAGAAATCAATGGAAGTCTCAGCTCA	954
Oy	1302	CCGATGCCCGGGGTGTGCTCCCTCTGTACCCACCGCAGCGGCGCTTGGAAATGGCCAGC	1361
Db	955	C-----GGAGGGCGGCTCCCGGCTGCTGCGGCGGCGGCGGCATTTGAGATGGCCAGC	1005
Oy	1362	ACAGCCAGAGATCCCGTAGCGTCAATGAGACCTTCCAGCGGTCTGTCTCCAGCCCTTAA	1421
Db	1006	ACGGCCAGAGGCTCTCGGTCCATCAGCGGTCTCTCAGGCTTTCACACCAAGCCACATCA	1065
Oy	1422	GGAGCCCAAGGAGTCCGGTCTTTCTCTTTCAACGGAGCGGGGGGCTGGAATGAGCTC	1481
Db	1066	GGAGCCCGGGGT-----	1078
Oy	1482	GAGGCGGGGCTCCCGGACTTCCAAAAACGACGAGCTTCTCGGGGGCCCAAGGGGTG	1541
Db	1079	-----	1078
Oy	1542	GGGGCGCGGGGAGAGCCCGCGCCCGCCAGATGCGGCTCCACACCCCTGCGGGGGCCCC	1601
Db	1079	-----GACCCCTCACCCCTTCACCAAGGGGAGTGCCTCTCC-----	1115
Oy	1602	CAGGCTCCCGCGCTCTGTGGCGGAGCCCTTGTGACTCGCTGTGCACAGCCCGCGG	1661
Db	1116	-----ACCCCAAGGGAGACGCTGCACACGCGCAAGG	1149
Oy	1662	CCAGTCCACCGGAGCCCGGGGACAACACACCCCAAGCCCGCGGCGGTGGCTCGGGG	1721

Db	1150	AGAGCCCGGCTGGCCAGGCGCCCAACCCACGAGCCCGCTCCAGGCCCC-----AGCGTCCGAG	1203
Qy	1722	GAGCGCCGCTGGAGGAGTGTCTCAACTCCATTCGGAACAAGATTCTCGGGTCCCGCTGCT	1781
Db	1204	GGGTGCTCTGGAGGGCGGGGCTCAACTCCTATCAGAAAGAGCTTCTGGGCTCACCCGCT	1263
Qy	1782	TTTCACCGCGCGCAAGATGCGAGGTCCTTACCGCTGAGAGATGATCCAGTCTTGACGCGCAGT	1841
Db	1264	TCCACCGCGCGAAACTGCAAGTTCGAGCGCCGGAGAGATGTCCAACTGACACCGAGT	1323
Qy	1842	CCTTCCCGGAGCTGGCAAAACGCTCTGTTCGGAGACTTATCTCTTGGACAAAGAG	1901
Db	1324	CGTCCCGAGAGCTGGCGAAGAGTCTGTGTTGGAGACTTATCAGCTGGGAGAGAGAG	1383
Qy	1902	AACAAATATTTCTCTGCTTAAAGGCAAACTCTTAGAGCATCAAAACGACATCTGCC	1961
Db	1384	AGCAATATTTCTCTGTGTCATCAAAAGCAAACTCTAGTTCATCAAGGCTGCAATCTGTC	1443
Qy	1962	ATGCGTTTCTGTCATGCCCGCAGCTGAGTGCACAGTGTCTGTACACAGCACAGCTTCAGG	2021
Db	1444	AGCGCTTCTGTGTCATTTCCCACTCTCAGCGACAGAGGTATCTCCCAAGCAGCTTCGGG	1503
Qy	2022	CCGAGTACAAAGGCCAGTGGCGGCGCCCTCGTTCGTTCCAAAAGCCGCTCCGCTTCCAAGTGG	2081
Db	1504	CCGAGTACAAAGGCCACGGGGGGGGCCAGCGGTGTCCAGAAAGCCGGTCAAGTTCAGGTTG	1563
Qy	2082	ACATCAGTCTCTTAGAGGTCAGAGCCCTCCCGCGAGCGGAGCGGACGCGGAGGTGGTG	2141
Db	1564	AAATCAGCTTACACGAGAGGT-----GGGAGGCGCGAGAAAGGAGAACG	1605
Qy	2142	GCATCTACGCTCAGCTTCACTCTCATCTCGGTGGTCCAGCGCGTGGTTCAGACGAGTGG	2201
Db	1606	GCATCTACGCTCAGCTTCACTCTCTGCTCTCAGGCGCCAGCGGTGCTTCAAGAGAGGTGG	1665
Qy	2202	TGGAGACCATTCAGGCGACAGCTCTGTAGACATCATGACCAGCCCTCCGCTGCAGGCCCTGG	2261
Db	1666	TGGAGACCATTCAGGCGCCAGCTGTGTGAGACACACAGACCCGCGCTCGGCGCCAGCACTTGT	1725
Qy	2262	CAGA	2265
Db	1726	CAGA	1729

AC020922 RESULT 11  
LOCUS AC020922  
DEFINITION Homo sapiens chromosome 19, clone CTD-2105E13, complete sequence.  
AC020922  
AC020922  
AC020922.8 GI:22038504  
HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 134792)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
Direct Submission  
Unpublished  
2 (bases 1 to 134792)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 134792)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
Direct Submission  
Submitted (01-AUG-2002) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On Aug 1, 2002 this sequence version replaced gi:15022017.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-shgc.stanford.edu

Quality: Phrap Quality >=40 99.4% of Sequence;  
Estimated Total Number of Errors is 0.4.  
NOTE: Shatter Libraries failed to resolve dinucleotide repeat  
region 104785-104880. Unsure number of repeat copies  
104785-104880. Forced join 104826.

FEATURES	Location/Qualifiers
source	1. .134792

misc\_feature

BASE COUNT	30161	a	37447	c	36890	g	30294	t
ORIGIN								

Query Match	22.8%;	Score 678.8;	DB 9;	Length 134792;
Best Local Similarity	99.4%;	Pred. No. 9.8e-118;		
Matches 692;	Conservative 0;	Mismatches 2;	Indels 2;	Gaps 1

[illegible]

RESULT 12  
AK074411

LOCUS	AK074411	3156 bp	mRNA	linear	PRI 15-FEB-2002
DEFINITION	Homo sapiens CDNA FLJ23831 fls, clone KAIM00425, highly similar to				
ACCESSION	Homo sapiens mRNA for putative serine/threonine protein kinase.				
VERSION	AK074411.1				
KEYWORDS	GI:18677005				
SOURCE	Oligo capping; fls (full insert sequence).				
	Homo sapiens ileal mucosa CDNA to mRNA, clone_1lb:KAlA				

## REFERENCE

## FEATURES

BASE COUNT	588 a	1075 c	957 g	536 t
ORIGIN				

Query Match	16.5%	Score 491.8;	DB 9;	Length 3156;
Best Local Similarity	61.9%	Pred, No. 1.6e-82;		
Matches 1075; Conservative	0;	Mismatches 372;	Indels 290;	Gaps 7

QY	529	CAGTAACTCCATCTGGCCACAGACAGACTAAAGCCCGAGAACCTGGTTTGGATGTGAGAAA	588
Db	3	CAACCTCTCTCCACGACACAGAGATCTGMAAACCTGAAAAACCTCTGCTGTGACGAGAAAGAA	62
QY	589	CAACATCCGATTCGACAGACTTCGGCAATGGCGTCCCTCAGGTGGGGGACACGCTCTCGA	648
Db	63	CAACATCCGATTCGACAGACTTTGGCAATGGCGTCCCTCAGGTGGGGGACACGCTGTGGA	122
QY	649	GACCAGCTGGGGGTCGCCCATTTATGCGTGCACAGAGTGATTAAGGGGGGAAAAATTGA	708
Db	123	GACCAGCTGGGGGTCCCCCACAATAGCGCTCCCGAGGTGATCCGGGGGGGAAAGTATGA	182
QY	709	TGGCGCGCGGGCAGACATGTGGAGCTGTGAGATCATCTCTTGGCCCTGCTCGTGGGGGC	768
Db	183	CGCGCGGAAAGCGGACGCTGTGGAGACTCGCGCGTCATCTGTTGGCCCTGCTGGTGGGGGC	242
QY	769	TCGCGCTTTGATGACAGCAACAACCTCCGCGCAGCTGTGTGGAAAGGTGAAGAGGGGCGCTT	828
Db	243	TCGCGCTTGTGAGG--	256
QY	829	CCACATGCCCCCACTTCAATTCTCTCAGATTGCCAGAGCGCTCTCTGAGGGGAATGACGAAT	888
Db	257	-----ATGATCGAGGT	267

QY 889 GAGAGCCGAAAAAGGCTCAGTGTGAGCAATTCAGAAACATCTTGTACTAGGGG 948  
 Db 268 GAGAGCCGCGAGCGCGCTCAGCTAGAGCACATTCAGAAACATATGTATAGGGG 327  
 QY 949 GAAACAGAGACGAGCCGCTGAGCCAGCCCTGAGCCGCGGAGGAGCATGCGAG 1008  
 Db 328 CAAAGATGAGCCGAGAC-----AGAGCAGCCATTCCTGCAAGGAGGAGATCCGCTC 381  
 QY 1009 CCGGCAATCCAGAGAGCTGAGACCCGAGCCTCTAGAGAGCATGGCATGAGGGCTG 1068  
 Db 382 GCGGCCAGGCTGAGAGACATGACCCCGAGCTGTGACAGCATGACTACTGTGGCTG 441  
 QY 1069 CTTCAGAGAGCCGAGAGAGCTGATGCGAGCTGCGAGTGAAGAGAGAGAGAGAG 1128  
 Db 442 CTTCAGAGAGCCGAGAGAGCTGATGCGAGCTGCGAGTGAAGAGAGAGAGAGAG 501  
 QY 1129 GATGATATATATCTGCTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1188  
 Db 502 GATGATATATCTGCTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 561  
 QY 1189 CCGGCTCCCGGAG 1248  
 Db 562 CCGGCTCCCGGAG 621  
 QY 1249 CCGGCTCCCGGAG 1308  
 Db 622 CCGGCTCCCGGAG 675  
 QY 1309 CCGGCTCCCGGAG 1368  
 Db 676 CCGGCTCCCGGAG 732  
 QY 1369 GAGATCCCGGAG 1428  
 Db 733 GAGATCCCGGAG 792  
 QY 1429 AAGAGATCCCGGAG 1488  
 Db 793 CCGGCTCCCGGAG 797  
 QY 1489 GAGATCCCGGAG 1548  
 Db 798 CCGGCTCCCGGAG 797  
 QY 1549 CCGGCTCCCGGAG 1608  
 Db 798 CCGGCTCCCGGAG 835  
 QY 1609 CCGGCTCCCGGAG 1668  
 Db 836 CCGGCTCCCGGAG 876  
 QY 1669 CCGGCTCCCGGAG 1728  
 Db 877 CCGGCTCCCGGAG 930  
 QY 1729 CCGGCTCCCGGAG 1788  
 Db 931 CCGGCTCCCGGAG 990  
 QY 1789 CCGGCTCCCGGAG 1848  
 Db 991 CCGGCTCCCGGAG 1050  
 QY 1849 CCGGCTCCCGGAG 1908  
 Db 1051 CCGGCTCCCGGAG 1110  
 QY 1909 CCGGCTCCCGGAG 1968  
 Db 1111 CCGGCTCCCGGAG 1170  
 QY 1969 CCGGCTCCCGGAG 2028

Db 1171 CCGTGTGATTCAGAGTGTGAGCAATTCAGAAACATCTTGTACTAGGGG 1230  
 QY 2029 CAAAGGAG 2088  
 Db 1231 CAAAGGAG 1290  
 QY 2089 CCGTGTGATTCAGAGTGTGAGCAATTCAGAAACATCTTGTACTAGGGG 2148  
 Db 1291 CCGTGTGATTCAGAGTGTGAGCAATTCAGAAACATCTTGTACTAGGGG 1332  
 QY 2149 CCGTGTGATTCAGAGTGTGAGCAATTCAGAAACATCTTGTACTAGGGG 2208  
 Db 1333 CCGTGTGATTCAGAGTGTGAGCAATTCAGAAACATCTTGTACTAGGGG 1392  
 QY 2209 CCGTGTGATTCAGAGTGTGAGCAATTCAGAAACATCTTGTACTAGGGG 2265  
 Db 1393 CCGTGTGATTCAGAGTGTGAGCAATTCAGAAACATCTTGTACTAGGGG 1449

RESULT 13  
 AB014885  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Tunicata; Ascidacea; Stolidobranchia; Pyuridae; Halocynthia.  
 1 (sites)  
 Sasakura, Y., Ogasawara, M. and Makabe, K.W.  
 Maternally localized RNA encoding a serine/threonine protein kinase in the ascidian, *Halocynthia roretzi*  
 Mech. Dev. 76 (1-2), 161-163 (1998)  
 9840280  
 2 (bases 1 to 3933)  
 Sasakura, Y., Ogasawara, M. and Makabe, K.W.  
 Direct Submission  
 Submitted (26-MAY-1998) Yasunori Sasakura, Kyoto University, Department of Zoology, Graduate School of Science, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan (E-mail: sasakura@ascidian.zool.kyoto-u.ac.jp, Tel: +81-75-753-4095, Fax: +81-75-705-1113)  
 Location/Qualifiers  
 1. 3933  
 /organism="Halocynthia roretzi"  
 /db\_xref="taxon:7729"  
 /tissue\_type="egg"  
 /dev\_stage="early embryonic"  
 /tissue\_1lb="fertilized egg"  
 1. 3933  
 /gene="HrPOK-1"  
 166. 2262  
 /gene="HrPOK-1"  
 /codon\_start=1  
 /protein\_id="BAA28663.1"  
 /db\_xref="GI:3172111"  
 /translation="MSNAPOGVOVGYRKYLEKLGKGTGLVKGVMGTGKVAVKI  
 VNRKLSSEYINKEVERIATIMKLEHHHIIIGHYVNRKLYLILEVSGELPDL  
 VOKRLIPREARFRFOITISAVDYCHNHNTCHRLKRENLIDKKNITKPADDFGMSL  
 OPEGLLETSCGSPHYATKINGERYDGRADWSCVILFALLVGLPDDDLRLQ  
 LLEKVRGVYHIFPVPDAONLRLGMLIDVAPDKRLIQVYLQHPMPGNSVEGLV  
 VEPDPAVYIDVCLPEESVDPLVLSMTSLGCFKKEKILKMLITEQWTEKVVY  
 MLRRKRYSPFDODADSLPCKHPDARKRYDSSLSNGDMCVNPIDPKRVAE  
 SLCTDSSPILSRKSTEHORSOSTLGSNSRLVONISPOTKAESKRIINGPYVR  
 GTTCSSNPQINTPASPNPWRRLASLKTTPGSPRFRHKKQAPSESEVENQNS  
 SSELKRSWFGNFWMSRYSTEHDELPLVALYKRLINSVSELVAFSLIPNTHS  
 WSPTRKCDYRSSTSTVFHORSIKFOVDIIQHSILDOENGRKPSQVGSFTI  
 AFSLISGPIRKYRVELLOMOMSVIAVONONKTSGEFSAPSHIIOCLKIVNMLHOD  
 NSYMAINPNPHKSKVYK"

polya\_site

3933



```

Db      322 TCAGATGATGCAACAAAGAAAGAACTGATGATGTTTACAAAGGTCGAAAGAGAAA 381
Oy      330 TCGCCATTCCTGAGATCATGACACACACATGCTCTCAAGCTCCAGACATCTACAGAGA 389
Db      382 TTGCATTTATGAAGTATGATAGACATCCACATGCTCTCAATTTTACGATCTACAGAAA 441
Oy      390 ACAAGAAATATTTGATGATGCTGTTCTGAGACAGCTCTCGGGGGGAGAGCTATTTGACATAC 449
Db      442 ATTAGAAATATTTATTTCTACTATGACATGTTAGCGGGGAGAGCTGTTGACATACC 501
Oy      450 TGTGAAGAAGGGGAGACATGACCCCAAGAGAGGCCCCAAAGTTCTTCCGACAGATGCT 509
Db      502 TTGTGGGAAAGGAGCGCTGATGTCGAAAGAGCGCAAGATTTTTCGTAATCAATCATCT 561
Oy      510 CTGCGCTGACCTCTGCGACAGCTACTCTCCATCTCCACAGACCTAAAGCCGAGAAC 569
Db      562 CTGCGCTGACCTTTTTCGCACTGCTATCAATTTCTCATCGAGATTTAAACCGGAAACT 621
Oy      570 TCGTTTGGATGAGAAACACACATCCGATTCGAGACTTCGCGCATGGCTCCCTGACAG 629
Db      622 TATTACTAGACAGAGAAATATTAAGTAGACAGATTTTGAATGCTAGTCTGACAG 681
Oy      630 TCGGGGACAGCCCTCTGAGACACAGCTGCGGCTCCCTCATATGCGTCCAGAGAGTGA 689
Db      682 TGGAAAGGTCGATGTTGAGACAGAGTGTGATCTCCATATGATGATGATCCAGAGTTA 741
Oy      690 TTAAAGGGGAAATATATGATGCGCCGCGGACAGATGTTGAGCTGTGATCATCTCT 749
Db      742 TTGCGGGTGAAGATACAGAGCTGCGAAAGACAGAGCTTTGAGAGTTGCGCTTATCCCT 801
Oy      750 TCGCCCTGCTCTGTTGGGGGCTCTGCGCTTTGATGACAGACACATCCGCGACGCTGAG 809
Db      802 ACCTATATATGATGTTGGCTCTTCCATTTGATGATGACAAATCTCGAAATCTTCTGAAA 861
Oy      810 AGGTGAAGAGGGGCTCTGATGATGCGCCCTCATCTCTCCAGATTCGACAGCCCTCC 869
Db      862 AAGTGAAGCTGAGATTTTTCATATCCACATTTGCTGATGCTGATGCTGAGAGTTTAC 921
Oy      870 TGAAGGGAATGATGATGAGAGGCGGCAAAAGAGCTGATCTGAGAGCAATTCAGAAAC 929
Db      922 TGGGGGAATGATGATGAGAGTGAAGCCGGGAAAGATATTCCTGAGCCAGCTTCAAC 981
Oy      930 ATCTTGTGATCTTACAGCGGGAACAGACAGCCGCTGCTGAGAGCAGCCCTGGCC 989
Db      982 ATCTATG---GATATCGGAAACAAAGGAGATCTGAAATTTGAGAGCTTCCATATGTCAC 1038
Oy      990 GCGGGGATGAGCGGAGCGCTGCTGATGCAAGGAGAGAGCTGAGCCGAGCGCTGAGAGA 1049
Db      1039 AAGTTGTGACAGAGCGACGTAATACCTGCTGAAAGATTCGATGATCGGATGCTTGAAGC 1098
Oy      1050 GCATGCGATCACTGGGCTGCTTACAGGACCGGAGAGGCTGCTGATGCGAGCTGCGCAGTG 1109
Db      1099 ATATGAACTGCTTATGATGTTTCAAGACAAAGAAAGTATGATTAATGAATATATGTCGC 1158
Oy      1110 AGAGAGGAAACCAAGAAAGATGATATATTTATCTGCTTTTGGATGCGAAGGAGCGCTAC 1169
Db      1159 CAAGAAACCAACACTGAAAGATGATATCTTCTCTCTTCTGTAAGAAACGACGCGCC 1218
Oy      1170 CCAGCTGTGAGACCA 1185
Db      1219 CTGACACAGGAAGACGA 1234

```

```

RESULT 15
LOCUS      AY060288          5609 bp      mRNA      linear      INV 05-NOV-2001
DEFINITION Drosophila melanogaster GH13047 full length cDNA.
ACCESSION  AY060288
VERSION     AY060288.1 GI:16648123
KEYWORDS   FLI_CDNA.
SOURCE      Drosophila melanogaster.
ORGANISM   Drosophila melanogaster

```

## REFERENCE

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 5609)  
Stapleton, M., Broksstein, P., Hong, L., Agbayani, A., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guatin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J., Nuno, J., Pacle, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M. and Celniker, S.

## TITLE

Submitted (29-Oct-2001) Berkeley Drosophila genome project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA

## COMMENT

Sequence submitted by:  
Berkeley Drosophila genome project  
Lawrence Berkeley National Laboratory  
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our web site (<http://fruitfly.berkeley.edu>) or send email to [cdna@fruitfly.berkeley.edu](mailto:cdna@fruitfly.berkeley.edu).

## FEATURES

Location/Qualifiers

```

1..5609
/organism="Drosophila melanogaster"
/strain="Y: cn bw sp"
/db_xref="taxon:7227"
/map="72A5-72B1"
1..3609
/gene="CG6114"
/note="alignment with genomic scaffold AE003529"
/db_xref="FlyBase:FBgn0036544"
372..2377
/gene="CG6114"
/note="Longest ORF"
/codon_start=1
/product="GH13047p"
/protein_id="AAL25327.1"
/db_xref="GI:16648124"
/db_xref="FlyBase:FBgn0036544"

```

## CDS

```

1..5609
/translation="MASLOPAGSMLETCGSPHYACEVIRGEYDGRKADWSCGYI
LVALVGALEFDNDLRLLEKVRGVFHPVPPQSLRGMIVNDRRLTLE
INHPVTVAGKGELELPMEEVVOHVPTAVDPVLAICSGCEKREKRLIO
ELISSHNTKEVYFLLEKRRRPALEDDDELAOKRSELDVADPPKRLDTCRIN
TNAPSGOISEGSLVPLPRAOAFRRFSYSTRNORRSTPTVTSVRSVSPYRCNS
PMSSAOQANALISRPSPACGTRSHYSGDPRGCHSSVSTPSSHOKSTIEGVVYV
REPRIRDSLRDERGGSPRDGDCITPGSGFGNSGTSASPSVHHRANSGPTIA
ISMFDPSNSVNPNGSPPMNNSSPMPSPGCTGGQMLKTRLTNIKNSFLSPR
HRKKQVSADEVHLIPRESSPELTKRSFNGMLITTEKDEFTLLVKGPIAVKXHLI
AFLSMAELSHSVSPTSFVREYKRNKGPMVQHRVFOVDISAICOGDIADMLFAL
TFLLSGTRRRRICEHIOGSKRFPDPSSPTVSVTOAESSSCGSVSERL
SYKROVLENDMDNDISFYSKSGNRRASTNNNNANVDIPGPIAVRNSNETAHER
NRELQSERQNTVSGSHIA"

```

## BASE COUNT

1803 a 1224 c 1164 g 1418 t

## ORIGIN

Query Match 12.5%; Score 373.6; DB 3; Length 5609;  
Best Local Similarity 65.5%; Pred. No. 3, 2e-60;

Matches 563; Conservative 0; Mismatches 294; Indels 3; Gaps 1;

```

Oy      344 CTGATGACACCCACATGCTCCAGCTCCAGACGCTTACGAGACAAAGAAATATTTG 403
Db      2 CTATGATCATCCACAGCTCTGCGCTGAGCGATGTTGTCAGAGAACAGATATTTG 61

```

OY 404 TACCTGTTCTGAGACGCTCTGGGGGTGAGCTATTGCACTAGTGTAAAGAGGG 463  
DB 62 TATTGATATTGAGCATGATCCGGGGAGAGCTTTGATTAATCTGTGTAAGAGGGT 121  
OY 464 AGACTAGACCCCAAGAGAGGCGCAAAATTTCTCGCCAGATTTGTCTGCGTGAATTC 523  
DB 122 CGATTGACGCGGAGAGGCGCGCAAGTTCTTCAGGCAATCATCTCCGCTGATTTTC 181  
OY 524 TGCACAGCTACTCCATCTCCACAGAGACTTAAAGCCCGAAGCTGCTTTGGATGAG 583  
DB 182 TGCCACTGCGATTCGATTTGCACTCGGACTTGAAGCCGGAATCTGTCTGTGAGAG 241  
OY 584 AAAACACATCCGATTTGAGACTTGGCATGCGCTCCCTGAGTGGGGGACAGCCTC 643  
DB 242 AAGAAATTAACATTAAGATAGCGGACTTTGGATGAGCTTCCCTGACAGCAGCTGGCAGCATG 301  
OY 644 CTGGAGACGCTGCGGGGTCCCGCATTAATGCTGTCCAGAGGTATTAAGGGGAAAA 703  
DB 302 TTGGAGACCTCTGCGGCGACGCCACTACGCTGTCCAGAGGTATACGGGCGAGAG 361  
OY 704 TATGATGCGCGCGGCGAGACATGTGAGCTGTGAGTCACTCTCCGCGCTGCTGTG 763  
DB 362 TACGATGCGCGGAGAGCGGATGTGTCTGTGTGGGTCACTCTATGCTCTCTGTG 421  
OY 764 GGGGCTGTGCTTGTATGACGAACTTCGCGAGCTGCTGAGAAAGTGAACGGGGC 823  
DB 422 GGTGGCTGCTCTGCGAGCAGCAACTTGGCAGCTGCTGAGAAAGTCAACGGGGC 481  
OY 824 GTCTCCACATGCGCCCATCTCATCTCCAGATTTGCCAGAGCTCTCTGAGGGAAATGTC 883  
DB 482 GTCTTTCATATACGCACTTTGTGCGCGGACTGCCAGAGTCTGTGCGGGCATGATT 541  
OY 884 GAAGTGGACCCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTTGTGATCTA 943  
DB 542 GAGTCAATCCGAGACGGGGGTCAAGCTGGCTGAATCAACGGGATCCGTGGGTACA 601  
OY 944 GCGGGAAAAACGAGCCAAACCGTCTGTGAGCCAGCCCTGCGCGGAGTACCATG 1003  
DB 602 GCTGCGCGGAAAAAGG--GAGCTGAGCTGAGCTGCAATGATGAGGTGTGACAGACA 658  
OY 1004 CGAGACCTGCAATCCAAAGAGAGTGAACCGGAGCTGCTAGAGCATGGCATCACTG 1063  
DB 659 CACGTTATTTCCACAGCCACCGGGTGTGATCCGATGTGTTGAACGGATTTGCTGCTG 718  
OY 1064 GCGTCTTCAGGAGCGCGAGAGGCTGCATCGGAGCTGCGAGTGGAGAGAGAACCAA 1123  
DB 719 GCGTCTTCAGAGAGAGAACTCATCCAGAGACTGCTCAGTTCAAGTCAAAATACG 778  
OY 1124 GAAAAGATGATATTATTGCTTTTGGATCGAGAGAGCGGTATCCAGCTGTGAGAC 1183  
DB 779 GAGAAGGTTATATTATTTCTGTGTGAGAGCGCAAGCAAGACGACCTCGCTGAGAGAT 838  
OY 1184 CAGGACCTGCTCCCGGAA 1203  
DB 839 GATGATGAGATTGCGCAAAA 858

Search completed: April 21, 2003, 23:04:02  
Job time : 5185.18 secs



XX 14-JUN-2001: 2001WO-US19444.  
 PF 15-JUN-2000: 2000US-212073P.  
 PR 23-JUN-2000: 2000US-213467P.  
 PR 30-JUN-2000: 2000US-215651P.  
 PR 07-JUL-2000: 2000US-216605P.  
 PR 13-JUL-2000: 2000US-218372P.  
 PR 25-AUG-2000: 2000US-228056P.  
 XX (INCYTE GENOMICS INC.)  
 PA Yue H, Lal P, Bandman O, Borowsky M, Au-Young J, Lu Y;  
 PI Ganhi AR, Tibboley CM, Walla NK, Yao MG, Lu DM, Greenwald SR;  
 PI Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YT;  
 PI Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;  
 PI Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JL, Ding L;  
 PI Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;  
 XX WPI: 2002:090307/12.  
 DR P-PSDB: ABE16271.  
 DR  
 PR New polypeptides, useful for diagnosing, treating or preventing  
 PI disorders of growth and development, cardiovascular and lipid, and  
 PI diseases such as cancer, comprise human kinase polypeptides -  
 XX  
 PS Claim 5; Page 188-189; 197pp; English.  
 XX  
 CC The invention relates to human kinase PKIN proteins and their  
 CC corresponding cDNAs. A composition containing PKIN agonist is useful for  
 CC treating a disease or condition associated with decreased expression of  
 CC PKIN and a composition comprising PKIN antagonist is useful for treating  
 CC a disease or condition associated with overexpression of PKIN. The  
 CC disorders include cancer (leukemia, adenocarcinoma, lymphoma, melanoma,  
 CC myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder  
 CC (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,  
 CC atherosclerosis, anemia, allergies, adult respiratory distress syndrome,  
 CC autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes  
 CC mellitus, multiple sclerosis, good pasture's syndrome, Graves' disease,  
 CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,  
 CC rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,  
 CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)  
 CC growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,  
 CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio  
 CC vascular disease (arteriovenous fistula, hypertension, vasculitis;  
 CC aneurysms, congestive heart failure, angina pectoris, myocarditis,  
 CC ischemic heart disease, chronic bronchitis, lung tumours); lipid  
 CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease); lipid  
 CC hypocholesterolemia, obesity). PKIN DNA is useful for assessing  
 CC toxicity of a test compound and in gene therapy. The present sequence  
 CC is human PKIN-17 cDNA.  
 XX  
 SQ Sequence 2897 BP: 599 A; 921 C; 877 G; 500 T; 0 other;  
 Query Match 91.3%; Score 2722.2; DB 24; Length 2897;  
 Best Local Similarity 99.1%; Pred. No. 0;  
 Matches 2747; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

QY 404 TACCTGTTCTGGAGCAGCTCTGGGGGTGAGTATTCAGTACCTGTTAAAGAGGG 463  
 DB 367 TACCTGTTCTGGAGCAGCTCTGGGGGTGAGTATTCAGTACCTGTTAAAGAGGG 426  
 QY 464 AGACTGAGCCCAAGAGAGGCCCAAGATTCTTCGCCAGATTGTCGTGGCTGACCTTC 523  
 DB 427 AGACTGAGCCCAAGAGAGGCCCAAGATTCTTCGCCAGATTGTCGTGGCTGACCTTC 486  
 QY 524 TGGCAGAGTCTACCTCCATCTGCGCAGAGACTTAAAGCCCAAGACCTGTTTGGATGAG 583  
 DB 487 TGGCAGAGTCTACCTCCATCTGCGCAGAGACTTAAAGCCCAAGACCTGTTTGGATGAG 546  
 QY 584 AAAAACACATCCGATTCAGACTTCGAGATGAGCTGCTCCCTGAGAGTGGGAGACGCTC 643  
 DB 547 AAAAACACATCCGATTCAGACTTCGAGATGAGCTGCTCCCTGAGAGTGGGAGACGCTC 606  
 QY 644 CTGAGAGCAGCTGGGGGTCCCAATATGCTGTCAGAGAGTGAATTAAGGGGAAAAA 703  
 DB 607 CTGAGAGCAGCTGGGGGTCCCAATATGCTGTCAGAGAGTGAATTAAGGGGAAAAA 666  
 QY 704 TATGATGGCCGGGCGGAGAGAGTGGAGCTGAGAGTATCTCTTCCCTGCTGCTG 763  
 DB 667 TATGATGGCCGGGCGGAGAGAGTGGAGCTGAGAGTATCTCTTCCCTGCTGCTG 726  
 QY 764 GGGGCTTGGCTTTGATGAGCAGCAACTCCGACAGCTGCTGAGAGAGTGAAGGGG 823  
 DB 727 GGGGCTTGGCTTTGATGAGCAGCAACTCCGACAGCTGCTGAGAGAGTGAAGGGG 786  
 QY 824 GCTTTCACATGCCCCACTTCAATCTCCAGATTCGCAAGAGCTCTGAGAGGAAATGATC 883  
 DB 787 GCTTTCACATGCCCCACTTCAATCTCCAGATTCGCAAGAGCTCTGAGAGGAAATGATC 846  
 QY 884 GAAGTGAAGCCCGAAAAAGCTCAGTCTGAGCAATTCAGAAATCCTTGCTGCTA 943  
 DB 847 GAAGTGAAGCCCGAAAAAGCTCAGTCTGAGCAATTCAGAAATCCTTGCTGCTA 906  
 QY 944 GGGGGAAGAACAGAGCAGACCCGCTGAGAGCCAGCCCTGAGCCGCGGTGAGCATG 1003  
 DB 907 GGGGGAAGAACAGAGCAGACCCGCTGAGAGCCAGCCCTGAGCCGCGGTGAGCATG 966  
 QY 1004 CGGAGCTGCTCCATCCAGAGAGAGTGGAGCCCGAGCTCTGAGAGCATGCTACTG 1063  
 DB 967 CGGAGCTGCTCCATCCAGAGAGAGTGGAGCCCGAGCTCTGAGAGCATGCTACTG 1026  
 QY 1064 GGGTCTTCAGAGAGCCGAGAGAGCTGATCCGAGCTGCGAGTGAAGAGGAGAACAA 1123  
 DB 1027 GGGTCTTCAGAGAGCCGAGAGAGCTGATCCGAGCTGCGAGTGAAGAGGAGAACAA 1086  
 QY 1124 GAAAAATGATATATATCTGCTTTGATGCGAAGAGAGCGGTATCCAGCTGTGAGAGC 1183  
 DB 1087 GAAAAATGATATATATCTGCTTTGATGCGAAGAGAGCGGTATCCAGCTGTGAGAGC 1146  
 QY 1184 CAGGAGCTGCTCCCGGAGATATGTTGACCCCGGAGAGCGTGTGATTCCTCATG 1243  
 DB 1147 CAGGAGCTGCTCCCGGAGATATGTTGACCCCGGAGAGCGTGTGATTCCTCATG 1206  
 QY 1244 CTGAGCCCTCAGAGAGAGCGGAGAGCGAGAGCGAGTCAATGGAAGTCTCTGAGCATCAC 1303  
 DB 1207 CTGAGCCCTCAGAGAGAGCGGAGAGCGAGAGCGAGTCAATGGAAGTCTCTGAGCATCAC 1266  
 QY 1304 GATGCGGGGGTGGTGGCTCCCTGTACCCAGCGAGCGGGCTTGGAGATGGCCAGCAC 1363  
 DB 1267 GATGCGGGGGTGGTGGCTCCCTGTACCCAGCGAGCGGGCTTGGAGATGGCCAGCAC 1326  
 QY 1364 AGCCAGAGATCCGTAAGGTGATGAGTGAAGCTCCAGAGGCTCTCTCAAGCCTTAAC 1423  
 DB 1327 AGCCAGAGATCCGTAAGGTGATGAGTGAAGCTCTCTCAAGGCTCTCTCAAGCCTTAAC 1386  
 QY 1424 AGCCAGAGATCCGTAAGGTGATGAGTGAAGCTCTCTCAAGGCTCTCTCAAGGCTCTCA 1483  
 DB 1387 AGCCAGAGATCCGTAAGGTGATGAGTGAAGCTCTCTCAAGGCTCTCTCAAGGCTCTCA 1446

OY	1484	GACGGGGGCTCCCCACATCTCCAAAAAGCAGAGCTCTCTTCGGGGCCCAAGGGGTGG	1543
Db	1447	GCGGGGGGCTCCCCACATCTCCAAAAAGCAGAGAGCTCTCTTCGGGGCCCAAGGGGTGG	1506
OY	1544	GCGCGCGGGGAGCAGCCCCCGCCCTCAGTGCCTCGCTCAGACACCTTCGCCGGCCCCCA	1603
Db	1507	GCGCGCGGGGAGCAGCCCCCGCCCTCAGTGCCTCGCTCAGACACCTTCGCCGGCCCCCA	1566
OY	1604	GCGTCCCGCGCGCTCGCTGGGGGAGAGCCCGCTTGCAGTGCCTCTGTGACAGAGCCCGGGCC	1663
Db	1567	GCGTCCCGCGCGCTCGCTGGGGGAGAGCCCGCTTGCAGTGCCTCTGTGACAGAGCCCGGGCC	1626
OY	1664	AGTCCACCGGGAGACCCCGGGGAGCAACACCACCCCGCAGCCCGGGCGTGGCGGGGA	1723
Db	1627	AGTCCACCGGGAGACCCCGGGGAGCAACACCACCCCGCAGCCCGGGCGTGGCGGGGA	1686
OY	1724	GCCCGCTGGAGAGAGTGTCTCAACTTCATCCGCAACAGCTTCTGGGGTCCCTCGCTTT	1783
Db	1687	GCCCGCTGGAGAGAGTGTCTCAACTTCATCCGCAACAGCTTCTGGGGTCCCTCGCTTT	1746
OY	1784	CACCGCGCAAGATCAGGTCCCTTACCGCTGAGAGATGTCCAGTCTGAGCGCAGAGTCC	1843
Db	1747	CACCGCGCAAGATCAGAGTCCCTTACCGCTGAGAGATGTCCAGTCTGAGCGCAGAGTCC	1806
OY	1844	TCGCCGAGCTGGCAAAACGCTCTGTGGTTCGGGACTTCATCTCTTGGACAAAGAA	1903
Db	1807	TCGCCGAGAGTGGCAAAACGCTCTGTGGTTCGGGACTTCATCTCTTGGACAAAGAA	1866
OY	1904	CAAAATTTCTGTGTCTAAAGGCAAAACCTTCAGAGAGCTCAAGAGAGCATGTGCAT	1963
Db	1867	CAAAATTTCTGTGTCTAAAGGCAAAACCTTCAGAGAGCTCAAGAGAGCATGTGCAT	1926
OY	1964	GCGTTTGTGTGCATCCCGCAGCCTGTAGTCAACAGTGCCTGTCCACAGACACAGTTCAGGGCC	2023
Db	1927	GCGTTTGTGTGCATCCCGCAGCCTGTAGTCAACAGTGCCTGTCCACAGACACAGTTCAGGGCC	1986
OY	2024	GAGTACAAAGGCCAGTGGGGGGCCCTCGCTTTCAAAAGCCCGTCCGGCTTCAGAGTGGAC	2083
Db	1987	GAGTACAAAGGCCAGTGGGGGGCCCTCGCTTTCAAAAGCCCGTCCGGCTTCAGAGTGGAC	2046
OY	2084	ATCAGCTCCTCTGAGAGGTTCCAGAGCCCTCCCGCGAGAGGGAGCGGACGGAGGTGTGGC	2143
Db	2047	ATCAGCTCCTCTGAGAGGTTCCAGAGCCCTCCCGCGAGAGGGAGCGGAGGTGTGGC	2106
OY	2144	ATCTACTCCGTCACTTCACTCTCATCTGGGGTCCCAAGCCGTGGTTCAAGCGAGTGGTG	2203
Db	2107	ATCTACTCCGTCACTTCACTCTCATCTGGGGTCCCAAGCCGTGGTTCAAGCGAGTGGTG	2166
OY	2204	GAGACCATCAGGACAGTCTCGTGAAGCAGTCAATGACAGCCCTCCGTCAGAGCCCTGGCA	2263
Db	2167	GAGACCATCAGGACAGTCTCGTGAAGCAGTCAATGACAGCCCTCCGTCAGAGCCCTGGCA	2226
OY	2264	GACGAGAAAGACGGGGGCCAGACCCGGCGCTGTGGTCCCGACCCCGAAGCCTTGACAGCC	2323
Db	2227	GACGAGAAAGACGGGGGCCAGACCCGGCGCTGTGGTCCCGACCCCGAAGCCTTGACAGCC	2286
OY	2324	GCACCCGGCGGCCAGACCCAGAGGTGAGAGTCTCCCGCGAGAGGCCCGCCCGCAAGAGC	2383
Db	2287	GCACCCGGCGGCCAGACCCAGAGGTGAGAGTCTCCCGCGAGAGGCCCGCCCGCAAGAGC	2346
OY	2384	AAGAAAGCTCTGTGGCCACAAAGGGAGCCCTCTCGCTGAGCCCGCAGGGGGCGGGAGGGA	2443
Db	2347	AAGAAAGCTCTGTGGCCACAAAGGGAGCCCTCTCGCTGAGCCCGCAGGGGGCGGGAGGGA	2406
OY	2444	GGGGAGCCCCCTTCACCCCGCTTCCGTGGCCCGCCCAACTGTGAATCTTAAATAAGGCCCA	2503
Db	2407	GGGGAGCCCCCTTCACCCCGCTTCCGTGGCCCGCCCAACTGTGAATCTTAAATAAGGCCCA	2466
OY	2504	AGGAACATGTGGGAGGGGGGTGACAACAAAACCGGCTTGGCCCTCAGGGAGTGGGGCT	2563
Db	2467	AGGAACATGTGGGAGGGGGGTGACAACAAAACCGGCTTGGCCCTCAGGGAGTGGGGCT	2526
OY	2564	CCACAGCGCGTCCCAACTGTGGGGTGGTTCTAGGGACACAGGGGGCGGGGAGCTGTTTTC	2623

Dd	25 27	CCACAGGCGGTGCCCAACTGGGGTGGTTCTAGGGGAACGAGGGGCGGGGAGCTGTTTC	2586
Oy	26 24	TATTTTATTATTGATTAAATTATTTATTTTATTATTGATCAATCTCTCTCGGGGGTGCG	2683
Dd	25 87	TATTTTATTATTGATTAAATTATTTATTTTATTATTGATCAATCTCTCGCGGGGTGCG	2646
Oy	26 84	GTCGGGGAGGAGCAGCGGAGCTGGTTGGGGTGCGCTTAGCAGATCCGGACAGGAGCCCTCTGC	2743
Dd	26 47	GTCGGGGAGGAGCAGCGGAGCTGGTTGGGGTGCGCTTAGCAGATCCGGAGAACGGGCCCCCTCTGTC	2706
Oy	27 44	CCTGTGTGTCGCCCAACCCCCTCTTCCCGGGCCCTCTCCCTGGTCC-TCCCCCCACG	2802
Dd	27 07	CGTGtgcTcTcccccAacccccctTtcTccCGggcccccTtcTccctgTgcTtcTcccccAcag	2766
Oy	28 03	ACCCTCTGTACGGAATTGCTCTCCGGAAGAATTCTGGTTGCGGTATCCTGCGCTCGT	2862
Dd	27 67	ACCTTCCTGTACGGAATTGCTCTCCGGAAGAATTCTGTATAACCGGTATCTCGCTGCGT	2826
Oy	28 63	CCGTGCTCTGTATCCGCCGGCGGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGTA	2922
Dd	28 27	CCGTGCTCTGTATCCGCCGGCGGCAAAAAAAAAAACACAAACCAACAACAGAGGGCAC	2886
Oy	29 23	ATAATATAAAA 2933	
Dd	28 87	AACAAAAAAA 2897	

CC	AA06717	standard; CDNA; 2385 BP.
CC	AA06717	
CC	AA06717;	
CC	12-SEP-2001	(first entry)
CC	Polynucleotide sequence encoding human protein kinase #17.	
CC	Human; protein kinase; PTK; STK; cancer; cardiovascular disease;	
CC	metabolic disorder; immune related disease; neurological disorder;	
CC	neurodegenerative disorder; inflammatory disorder; infectious disease;	
CC	reproductive disorder; gene therapy; ss.	
CC	Homo sapiens.	
CC	WO200138503-A2.	
CC	31-MAY-2001.	
CC	22-NOV-2000; 2000MO-US32085.	
CC	24-NOV-1999; 99US-0167482.	
CC	(SDGE-) SUGEN INC.	
CC	Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;	
CC	Flanagan P, Clary D;	
CC	WPI; 2001-343950/36.	
CC	P-PSDB; AAU03517.	
CC	Nucleic acids encoding human kinase polypeptides, useful for preventing	
CC	diagnosing and/or treating e.g. cancer, immune, cardiovascular and	
CC	neutonal-associated diseases, and microbial infections -	
CC	Example 1; Figure 1; 433pp; English.	
CC	AA06701-AA06757 encode for novel human protein kinases #1-57. The	
CC	novel protein kinases have been identified as members of the tyrosine	
CC	or serine/threonine kinase (PTK and STK) families. The polynucleotides	
CC	encoding protein kinases and the polypeptides may be used in the	
CC	prevention, diagnosis and treatment of diseases associated with	
CC	inappropriate kinase expression. For example, they may be used to treat	

CC cancers (especially cancers of hematopoietic origin), cardiovascular  
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
 CC immune related diseases (e.g. rheumatoid arthritis), neurological  
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
 CC Additionally, polynucleotides encoding protein kinases may be  
 CC used for gene therapy and as DNA probes in diagnostic assays.  
 CC The protein kinase polypeptides may be used as antigens in the production  
 CC of antibodies against the protein kinases and in assays to identify  
 CC modulators of protein kinase expression and activity.  
 CC  
 XX  
 SQ Sequence 2385 BP; 494 A; 768 C; 731 G; 392 T; 0 other;

Query Match 75.8%; Score 2259; DB 22; Length 2385;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 CAGGCCAATATGTTGGGCTATGCGCTGAGAGAGAGCTGGGCAAGGACAGACAGG 223  
 DB 127 CAGGCCAATATGTTGGGCTATGCGCTGAGAGAGAGCTGGGCAAGGACAGACAGG 186  
 QY 224 CTGCTTAACTCGGGGTCACATGCTACAGGCTCAGAGGTCGCCATCAAGATGCTGAC 283  
 DB 187 CTGCTTAACTCGGGGTCACATGCTACAGGCTCAGAGGTCGCCATCAAGATGCTGAC 246  
 QY 284 CCGGAGAGCTGTGCGAGTGTGCTGTGAGAGAGGCGGAGATGCCATCTGAAG 343  
 DB 247 CCGGAGAGCTGTGCGAGTGTGCTGTGAGAGAGGCGGAGATGCCATCTCTGAAG 306  
 QY 344 CTCATGAGACCCACATGCTCTCAAGCTCAGAGCTACAGAGTACAGAGAAATATTG 403  
 DB 307 CTCATGAGACCCACATGCTCTCAAGCTCAGAGCTACAGAGAAATATTG 366  
 QY 404 TACCTGTTCTGAGAGCAGTCTGCGGGGTGAGCTATTCTACTACTCTGTAAGAGGG 463  
 DB 367 TACCTGTTCTGAGAGCAGTCTGCGGGGTGAGCTATTCTACTACTCTGTAAGAGGG 426  
 QY 464 AGACTGAGCCCAAGAGAGGCGGAAAGTTCTTCCGCAAGTTGTGTGCGTGGACTTC 523  
 DB 427 AGACTGAGCCCAAGAGAGGCGGAAAGTTCTTCCGCAAGTTGTGTGCGTGGACTTC 486  
 QY 524 TCCACAGCTACTCTCATCTGCGCACAGAGACTTAAGCGGAGAGCTTTGGATGAG 583  
 DB 487 TCCACAGCTACTCTCATCTGCGCACAGAGACTTAAGCGGAGAGCTTTGGATGAG 546  
 QY 584 AAAAACAACATCCGATTCAGACTTCGCGATGCGCTCCCTGAGAGTGGGAGAGCTTC 643  
 DB 547 AAAAACAACATCCGATTCAGACTTCGCGATGCGCTCCCTGAGAGTGGGAGAGCTTC 606  
 QY 644 CTGGAGACAGCTGCGGGTCCCGCCATTAATGCTGTCCAGAGTGAATTAAGGGGAAAA 703  
 DB 607 CTGGAGACAGCTGCGGGTCCCGCCATTAATGCTGTCCAGAGTGAATTAAGGGGAAAA 666  
 QY 704 TATGATGCGCGCGGAGACATGTTGAGACTGTGAGTCACTCTTCCGCTGCTGTG 763  
 DB 667 TATGATGCGCGCGGAGACATGTTGAGACTGTGAGTCACTCTTCCGCTGCTGTG 726  
 QY 764 GGGGCTGCTTTGATGAGACACAACCTCCGACAGCTGTGAGAGAGTGAACGGGGC 823  
 DB 727 GGGGCTGCTTTGATGAGACACAACCTCCGACAGCTGTGAGAGAGTGAACGGGGC 786  
 QY 824 GTCCTTCCACATGCGCCACTTCAATCTCTCCAGATGCGCAAGCTCTGAGGGAGATC 883  
 DB 787 GTCCTTCCACATGCGCCACTTCAATCTCTCCAGATGCGCAAGCTCTGAGGGAGATC 846  
 QY 884 GAATGAGAGCCGAGAAAAAGCTCAGTGTGAGCAAAATTCAGAAACATCTTGATCTA 943  
 DB 847 GAATGAGAGCCGAGAAAAAGCTCAGTGTGAGCAAAATTCAGAAACATCTTGATCTA 906  
 QY 944 GCGGAGAAACAGAGACCGAGACCTGCGCTGAGAGCCAGCCCTGGCGCGGTAGCCATG 1003  
 DB 907 GCGGAGAAACAGAGACCGAGACCTGCGCTGAGAGCCAGCCCTGGCGCGGTAGCCATG 966

QY 1004 CCGAGCTTCATCCAAAGAGAGTGAACCCCGAGCTCTAGAGACATGATGACTG 1063  
 DB 967 CCGAGCTTCATCCAAAGAGAGTGAACCCCGAGCTCTAGAGACATGATGACTG 1026  
 QY 1064 GCGCTTTAGAGAGCCCGAGAGCTGATCGGAGCTGCGAGTGAAGAGAGAACCA 1123  
 DB 1027 GCGCTTTAGAGAGCCCGAGAGCTGATCGGAGCTGCGAGTGAAGAGAGAACCA 1086  
 QY 1124 GAAAGATATATATATATGCTTTTGGATGCGAGAGAGGATATCCACTGTGAGAC 1183  
 DB 1087 GAAAGATATATATATATGCTTTTGGATGCGAGAGAGGATATCCACTGTGAGAC 1146  
 QY 1184 CAGGACCTCTCCCGGAGATGATGTTGACCCCGGAGAGGCTGATGCTCCATG 1243  
 DB 1147 CAGGACCTCTCCCGGAGATGATGTTGACCCCGGAGAGGCTGATGCTCCATG 1206  
 QY 1244 CTGAGCCGTACAGGAGAGCGGAGCCAGAGAGGATGATGAGAGTCTGAGCATCAC 1303  
 DB 1207 CTGAGCCGTACAGGAGAGCGGAGCCAGAGAGGATGATGAGAGTCTGAGCATCAC 1266  
 QY 1304 GATGCGGGGGTGGTGGCTCCCTGTAACCAACCGAGCGCTTGGAGATGCGCCAGAC 1363  
 DB 1267 GATGCGGGGGTGGTGGCTCCCTGTAACCAACCGAGCGCTTGGAGATGCGCCAGAC 1326  
 QY 1364 AGCCAGAGATCCCGTAGAGCTCAGTGAAGCTCCACAGGCTGTCTCCAGACCTTAAG 1423  
 DB 1327 AGCCAGAGATCCCGTAGAGCTCAGTGAAGCTCCACAGGCTGTCTCCAGACCTTAAG 1386  
 QY 1424 AGCCCAAGAGATCGGCTCTTTCTTTTCAACGAGAGCGGGGGCTGAGATGAGCTGA 1483  
 DB 1387 AGCCCAAGAGATCGGCTCTTTCTTTTCAACGAGAGCGGGGGCTGAGATGAGCTGA 1446  
 QY 1484 GCGGGGGCTCCCGCACTTCCAAAGCGAGAGCTCTTCTGGGGGCCCAAGGGGTGG 1543  
 DB 1447 GCGGGGGCTCCCGCACTTCCAAAGCGAGAGCTCTTCTGGGGGCCCAAGGGGTGG 1506  
 QY 1544 GCGCGGGGAGACAGCCCGCCCGCCAGTGCCTGTCACACCCCTGCGCGCCCA 1603  
 DB 1507 GCGCGGGGAGACAGCCCGCCCGCCAGTGCCTGTCACACCCCTGCGCGCCCA 1566  
 QY 1604 GGGCTCCCGGCTCTCTGCGGGAGACCCCTGTCAGTCCGCTGTGACAGCCCGGGCC 1663  
 DB 1567 GGGCTCCCGGCTCTCTGCGGGAGACCCCTGTCAGTCCGCTGTGACAGCCCGGGCC 1626  
 QY 1664 AGTCCACCGGGAGCCCGGGAGACAACACACCCAGCGCGGGGCGTGGGGGA 1723  
 DB 1627 AGTCCACCGGGAGCCCGGGAGACAACACACCCAGCGCGGGGCGTGGGGGA 1686  
 QY 1724 GCGGCTGAGAGAGTGTCTCAATCCATCGCAACAGCTTCTGCGCTCGCTT 1783  
 DB 1687 GCGGCTGAGAGAGTGTCTCAATCCATCGCAACAGCTTCTGCGCTCGCTT 1746  
 QY 1784 CACGGGGCAAGATGCAAGTCCCTACCGCTGAGAGATGCCAGTTCAGCCAGAGTCC 1843  
 DB 1747 CACGGGGCAAGATGCAAGTCCCTACCGCTGAGAGATGCCAGTTCAGCCAGAGTCC 1806  
 QY 1844 TCCCGGAGCTGGCAAAAGCTCTGTTGCGGAACTTCATCTCTTGAAGAAAGAA 1903  
 DB 1807 TCCCGGAGCTGGCAAAAGCTCTGTTGCGGAACTTCATCTCTTGAAGAAAGAA 1866  
 QY 1904 CAAATATCTCTGCTTAAGAGACAACCTTCAGACAGATCAAAAGCAGACATCTGTCAT 1963  
 DB 1867 CAAATATCTCTGCTTAAGAGACAACCTTCAGACAGATCAAAAGCAGACATCTGTCAT 1926  
 QY 1964 GCGTTCTGTGATGCCAGCTGAGTCAAGTGTGCTGACAGACAGCTTCAGGGCC 2023  
 DB 1927 GCGTTCTGTGATGCCAGCTGAGTCAAGTGTGCTGACAGACAGCTTCAGGGCC 1986  
 QY 2024 GAGTCAAGAGCGAGTGGGGCCCTCCGCTTCCAAAGAGCCGCTGCTCAAGGTGAGC 2083  
 DB 1987 GAGTCAAGAGCGAGTGGGGCCCTCCGCTTCCAAAGAGCCGCTGCTCAAGGTGAGC 2046





Matches 1477; Conservative 0; Mismatches 435; Indels 195; Gaps 6

[illegible]

Db 1702 TGCAGCCCTTCTGTGATTCACAGTCCAGTCCAGCCAGCCGTCATCTCCCAACGAGCTTCC 1761  
 QY 2019 GGGCGAGTACAAAGCCAGTGGCGGCCCTCCGCTTCCAAAAGCCGCTCCGCTTCAGG 2078  
 Db 1762 GGGCCAGATCAAGGGCCAGCGGGGGGGCCAGCCGCTGTCCAGAAAGCCGCTCAAGTTCAGG 1821  
 QY 2079 TGGACATCAGCTCCCTGCTGAGGCTCCAGAGCCCTCCCGCCAGCGGAGCGGAGAGGTG 2138  
 Db 1822 TTGATATCAGCTTACAGGAGGCT-----GGGGAGCGCCGAGAAAGAGA 1863  
 QY 2139 GTGGCATCTACTCCGTCACCTCTCATCTCGGGTCCAGCCGCTCGGTTCAAGCGAG 2198  
 Db 1864 AGGGCATCTACTCCGTCACCTCTCATCTCGGGTCCAGCCGCTCGGTTCAAGAGG 1923  
 QY 2199 TGGTGAGACATCATCAGGACGACCTCTCTAGACACTATGACACAGCCCTCCGTCAGAGCCC 2258  
 Db 1924 TGGTGAGACATCATCAGGACGACCTCTCTAGACACTATGACACAGCCCTCCGTCAGAGCCC 1983  
 QY 2259 TGGCAGA 2265  
 Db 1984 TGTTCAGA 1990

## RESULT 5

AAD34315  
 ID AAD34315 standard; cDNA; 2647 BP.

AAD34315;

16-JUL-2002 (first entry)

Human PKIN-18 cDNA.

Human; kinase; enzyme; PKIN-18 protein; immune system disorder; anaemia;  
 acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;  
 asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;  
 AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;  
 leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;  
 Down's syndrome; gene therapy; protein therapy; cytosolic; gene; ss.

Homo sapiens.

Key Location/Qualifiers  
 CDS 1..1995  
 /tag= a  
 /product= "Human PKIN-18 protein"

WO200218557-A2.

07-MAR-2002.

31-AUG-2001; 2001WO-US27219.

31-AUG-2000; 2000US-229873P.

08-SEP-2000; 2000US-231357P.

14-SEP-2000; 2000US-232654P.

22-SEP-2000; 2000US-234902P.

29-SEP-2000; 2000US-236495P.

06-OCT-2000; 2000US-238389P.

13-OCT-2000; 2000US-240542P.

(INCY-) INCYTE GENOMICS INC.

Bhandman O, Nguyen DB, Walla NK, Hafalia AFA, Yao MG, Gandhi AR;  
 Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;  
 Thornton M, Elliott VS, Lu Y, Ison CH, Au-Yang J, Tang YT;  
 Azimzai Y, Burdill JD, Marcus GA, Zingler KA, Lu DM, Lai PG;  
 Rankumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;  
 Burford N;

WPI; 2002-329769/36.

P-PSDB; AAE21723.

XX

PR New human kinases, useful for diagnosing, treating or preventing immune  
 system disorders (e.g. Crohn's disease), neurological disorders (e.g.  
 epilepsy), or cell proliferative disorders (e.g. cancers such as  
 leukemia or lymphoma)  
 PS Claim 97; Page 212-213; 218pp; English.

The present invention relates to human kinases (PKIN) and polynucleotides  
 encoding such proteins. PKIN sequences of the invention are useful for  
 diagnosing, treating or preventing disorders associated with aberrant  
 expression of PKIN, particularly immune system disorders (e.g. acquired  
 immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,  
 anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-  
 Tooth disease or seizures), cell proliferative disorders (e.g. cancers  
 such as adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma),  
 and developmental disorders (e.g. Down's syndrome). They are also used  
 in gene therapy and protein therapy. The present sequence is a cDNA  
 encoding human PKIN-18 protein.

Sequence 2647 BP; 525 A; 885 C; 789 G; 448 T; 0 other;

Query Match 32.1%; Score 955.8; DB 24; Length 2647;

Best Local Similarity 70.0%; Pred. No. 1.2e-177;

Matches 1473; Conservative 0; Mismatches 437; Indels 195; Gaps 6;

QY 173 TATGTGGGCCCCCTATCGGCTGAGAGAGACGCTGGCAAGAGACAGAGGCTGTTAA 232  
 Db 31 TATGTGGGCCCCCTATCGGCTGAGAGAGACGCTGGCAAGAGAGAGGCTGTTAA 90  
 QY 233 CTGGGGGTCACATCATCAGGAGGTCAGAAAGTCCCATCAAGTCTGAAACCGGAGAG 232  
 Db 91 CTGGGGGTCACATCATCAGGAGGTCAGAAAGTCCCATCAAGTCTGAAACCGTGAAG 150  
 QY 293 CTGTGGAAGTGGTGTGTATGAAGTGGAGCGGAGATGCCCTCGTGAAGCTCATCGAA 352  
 Db 131 CTCAGAGAGTGTGTGTATGAAGTGGAGCGGAGATGCCCTCGTGAAGCTCATCGAA 210  
 QY 353 CACCCACATGTCCTCAAGCTCAGCAGAGCTCTACGAGAACAGAAATATTTGTACTGTG 412  
 Db 211 CACCCACATGTCCTCAAGCTCAGCAGAGCTCTACGAGAACAGAAATATTTGTACTGTG 270  
 QY 413 CTGGAGACAGTCTCGGGGGGTGAGATTCGACTACTGCTGAAGAGGAGAGACTGACG 472  
 Db 271 CTAGAACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 330  
 QY 473 CCAAGAGAGGCGGAAAGTCTTCGCGCAGATGTGTGCGGTGAGTGTGAGTGTGAGTGTGAG 532  
 Db 331 CCTAAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 390  
 QY 533 TATCGATCTGCGACAGAGACTTAAAGCCGAGAACCTGTTTGGATGAGAAAAAAC 592  
 Db 391 CACTCATATGCGACAGAGAGTGTAAACCTGTTTGGATGAGAAAAAAC 450  
 QY 593 ATCCGATTCGAGACTTGGGAGTGTGCGAGTGTGCGAGTGTGCGAGTGTGCGAGTGTGCGAG 652  
 Db 451 ATCCGATTCGAGACTTGGGAGTGTGCGAGTGTGCGAGTGTGCGAGTGTGCGAGTGTGCGAG 510  
 QY 653 AGCTGCGGGTCCCGCCATATGCTGTCCAGAGAGTATTAAGGGGAGAAAAATATGATGGC 712  
 Db 511 AGCTGCGGGTCCCGCCATATGCTGTCCAGAGAGTATTAAGGGGAGAAAAATATGATGGC 570  
 QY 713 CGCGGGGAGAGATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 772  
 Db 571 CGGAAGGGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 630  
 QY 773 CCGTTTGAAGAGAACTCCGCGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 832  
 Db 631 CCGTTTGAAGAGAACTCCGCGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 690  
 QY 833 ATGCCCACTTATCTCTCAAGATGCGAGAGCTTCTGAGGGAAATGATCGAAGTGGAG 892  
 Db 691 ATGCCCACTTATCTCTCAAGATGCGAGAGCTTCTGAGGGAAATGATCGAAGTGGAG 750

QY 893 CCGGAAAAAGGCTCAGTCTGGAGCAATTCAGAAACATCTTGGTACTTACGGGGAAA 952  
 DB 751 GCCGACAGCCGCTCAGCTCAGAGCACATTCAGAAACATATGATATAGGGGCAAG 810  
 QY 953 CACGAGCCAGACCCGCTGGAGCCAGCCCTGGCCGGGTAGCCATGCGAGACTG 1012  
 DB 811 AATGAGCCGAGAAC-----AGAGCAGCCATCTCTCGAAGGTGCAAGTCCGCTGCTG 864  
 QY 1013 CCATTCAGAGAGCTGGAGCCGACCTCTAGAGAGCAGATGACATCTAGGCTGCTTC 1072  
 DB 865 CCGAGCTTGGAGAGATCAGACCCGACCTGCTGACACATCAGATCAGTGGGCTCTTC 924  
 QY 1073 AGGAGCGCGAGAGCTGAGAGTGGGAGTGGGAGTGGAGAGAGCAACCAAGAAAGATG 1132  
 DB 925 CGAGACCGCAACAGCTGCTCAGAGACCTGCTCGAGAGAGAGAAACCAAGAGAAATG 984  
 QY 1133 ATATATATATCTGCTTTTGGATCGAGAGAGCGGTATCCAGCTGTGAGAGCAGACCTG 1192  
 DB 985 ATTTACTTCTCTCTGAGAGCCGAAAGAAAGATACCCGAGCAGAGAGATGAGACCTG 1044  
 QY 1193 CCTCCCGGAAATGATGTTGACCCCGGAGAGCGTGTGATGTCATGCTGAGCCGT 1252  
 DB 1045 CCCCCCGGAAAGAGATAGACCTCCCGGAAAGCTGTGAGATCTCCCGATGCTGAACCG 1104  
 QY 1253 CAGCGGAAAGGCGGAGCAGAGCGGAAATGCAATGAGATCTGAGATCAGCATGCGGG 1312  
 DB 1105 CAGCGCAAGGCGGCGGCAAGAGCAAAATGAGAGTGTAGGTGAGTGC-----G 1155  
 QY 1313 GGTGGTGGCTCCCTGTACCCAGCCGAGCGGCTTGGAGATGGCCAGACAGACAGAGA 1372  
 DB 1156 GAGGCGGCTCCCGGTGCTGCGGGGGGCGCATGAGATGGCCAGACAGCGAGAG 1215  
 QY 1373 TCCCGTACGCTGAGTGGAGCCTTCCAGGGGCTGCTCCAGCCCTCTAAAGCCCAAG 1432  
 DB 1216 TCTGCTGCTACACAGCGGTGCTCTCAAGGCTTTCACAGCCCACTGACAGCC----- 1270  
 QY 1433 AGTCCGCTCTTCTCTTTCACCGAGCGGGGCTGAGATGAGAGCTGTGAGGCGGGG 1492  
 DB 1271 ----- 1270  
 QY 1493 TCCCCGATTCGAAAGCAGACGCTGCTTCTCGGGGCGCCAGGGGTGGGGCGCGG 1552  
 DB 1271 -----CCCGGGT 1277  
 QY 1553 GAGCAGCCCCCGCCCGAGTCCCGCTTCCACACCCCTGCGCCCGCCCGAGCTCCCG 1612  
 DB 1278 GACCCCTACCCCTCACCAGAGGGGAGTCCCTCC----- 1314  
 QY 1613 CGCTCTCTGGCGGGAGCCCTTGCACCTGCTTGCACAGCCCGCGGCACTCCACC 1672  
 DB 1315 -----ACCCCAAGGGGACCTGTTCACACGCCCAAGAGAGCCCGGCT 1359  
 QY 1673 GGCACCCCGGGGACACACACACCCCGAGCCCGGGGAGGGGAGCGCGCTG 1732  
 DB 1360 GGCACGCCCAACCCAGCCCGGCTCCAGCC-----AGCTGGGAGGGTGCCTG 1413  
 QY 1733 AGGAGCTGTCAACTCCATCCGACAGCTTCTGGCTCCCTGCTTTCACCGCGC 1792  
 DB 1414 AGGCGCGGCTCACTCATCAGAAACAGCTTCTGGGCTCACCCGCTTCCACCGCGG 1473  
 QY 1793 AAGATGAGAGTCCCTACCGCTGAGAGAGATGTCAGCTGAGCCAGAGTCTCCCGGAG 1852  
 DB 1474 AACTGTGAAGTTCGAGCGCCGAGAGATGTCACACGAGAGTCTCCAGAG 1533  
 QY 1853 CTGGCAAAACGCTCTGGTGGGAACTTCAATCTCTGAGCAAAAGAAATAATATTC 1912  
 DB 1534 CTGGCGAAGAGTCTGTTGGGAATCTTCACTGAGCTGGAGAGAGAGAGATCTTC 1593  
 QY 1913 CTGCTGTAAAGCAAACTTCAAGCATCAAGCAAGAGAGATGCTCATCTTCTG 1972  
 DB 1594 GTGCTCATCAAGCAAACTCTGAGCTTCATCAAGGCTGACATGCTGACCGCTTCTG 1653  
 QY 1973 TCGATCCCGAGCCCTGAGTCAAGATGCTGTCAAGAGCAGCTTCAAGGCGAGATCAAG 2032

DB 1654 TCGATTTCCACTTCCAGCAGCGTCAATCTCCAAAGAGCTTCCGGGCGAGTACAG 1713  
 QY 2033 GCCAGTGGCGGCCCCCTCGCTTCCAAAAGCCGCTCCGCTTCCAGGTGACATCAGCTCC 2092  
 DB 1714 GCCAGGGGGGGCCAGCGCTTCCAGAGAGCGGCTCAAGTTCAGATGATTCACCTAC 1773  
 QY 2093 TCTGAGGCTCCAGAGCCCTCCCGGAGCGGAGCGGAGCGGAGTGTGTCATCTCC 2152  
 DB 1774 ACGGAGGCT-----GGGAGGCGGAGAAAGAGAACGCGCATCTACTCC 1815  
 QY 2153 GTCACTTCACTCTCATCTGGGTCCAGCCGCTGCTCAAGCGAGTGTGAGACATC 2212  
 DB 1816 GTCACTTCACTCTCTCTCAGGCCCCAGCGCTGCTTCAAGAGGCTGTGAGACATC 1875  
 QY 2213 CAGGACAGCTCTGAGCACTCATGACACAGCCCTTCTGACAGCCCTGAGACAGAG 2272  
 DB 1876 CAGGCGCAGCTGTGAGACACAGACAGCCGCTGTGGGCCAGCACTTGTGACACACT 1935  
 QY 2273 AACGG 2277  
 DB 1936 AACGTG 1940

RESULT 6  
 ABA08296  
 ID ABA08296 standard; cDNA; 906 BP.  
 XX  
 AC ABA08296;  
 XX  
 DT 11-JUN-2002 (first entry)  
 XX  
 DE Human HRPOR-1 homologue-encoding cDNA, SEQ ID NO:72.  
 XX  
 KW Human: cytokine; cell proliferation; tissue growth; immunomodulator; growth factor;  
 KW haematopoiesis; regulation; tissue growth; immunomodulator; growth factor;  
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 KW chronic inflammatory condition; proliferative retinopathy;  
 KW atherosclerosis; coronary heart disease; arterial ischaemia;  
 KW bone disorder; osteoporosis; vascular growth disorder;  
 KW tissue regeneration; wound healing; infection; immune disorder;  
 KW cell culture; drug screening; gene therapy; antiinflammatory;  
 KW antiasthmatic; antiarthritic; haemostatic; antiatherosclerotic;  
 KW cytoskeletal; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
 KW antifungal; vulnery; antidiocier; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157188-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PE 05-FEB-2001; 2001WO-0503800.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 XX  
 PA (HYSE-) HYSO INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 DR WPI; 2001-45740/49.  
 DR P-PDB; ABB11052.  
 PT Human proteins and DNA encoding sequences useful for preventing,  
 PT treating or ameliorating a medical condition in a mammalian subject  
 PT e.g. arthritis and cancer -  
 XX  
 PS Claim 1; Page 364-365; 1963pp; English.  
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and

sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth factor activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a cDNA encoding a novel human polypeptide of the invention.

Sequence 906 BP; 205 A; 253 C; 267 G; 181 T; 0 other:

Query Match 22.1%; Score 658.8; DB 22; Length 906;  
Best Local Similarity 87.9%; Pred. No. 1.3e-119;  
Matches 786; Conservative 0; Mismatches 2; Indels 106; Gaps 2;

164 CAGGCCCAATATGTCGGCCCTATCGGCTGAGAGAACGCTGGGCAAGGACAGCAGG 223  
1 CAGGCCCAATATGTCGGCCCTATCGGCTGAGAGAACGCTGGGCAAGGACAGCAGG 60  
224 CTGCTTAACTCGGGGCTCACTGCATCAGGAGTGAAGGTCGCATCAAGATCGTGAAC 283  
61 CTGCTTAACTCGGGGCTCACTGCATCAGGAGTGAAGGTCGCATCAAGATCGTGAAC 120  
284 CCGGAGAACTGTCGAGTGGTGTCTGATGAAGGTGAGCGGAGATGCCATCT-GAA 342  
121 CCGGAGAACTGTCGAGTGGTGTCTGATGAAGGTGAGCGGAGATGCCATCTCTGAG 180  
343 GCTCATCGAACACCCACATCTCTCAAGCTCCAGAGCTTACAGAGAACAAATATTT 402  
181 GCTCATCGAACACCCACATCTCTCAAGCTCCAGAGCTTACAGAGAACAAATATTT 240  
403 ----- 402  
241 TCCCCCGGATGAATCAGATCAGTCCGTGATCTGGCGCAGTTTACACACGGGAA 300  
403 -----GTACCTGGTCTGGA 417  
301 GCTGTGCGCCCGGCTCATGAGACTGCTCTCAGGCTTTCACGATCGTCTGGA 360  
418 GCACGCTCTGGGGGGTGTGATGATGCTGATGATGATGATGATGATGATGATGATGAT 477  
361 GCACGCTCTGGGGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
478 GGAGGCCCGAAATGTTCTCGGCGAGATGTTGTCGCGTGAAGTTCCTCCAGAGTACTC 537  
|||||

421 GGAGGCCCGAAATGTTCTCGGCGAGATGTTGTCGCGTGAAGTTCCTCCAGAGTACTC 480  
538 CATCTCCACAGACACTTAAGCCCGAGAACCTGCTTTGGATGAGAAAAACAATCCG 597  
481 CATCTCCACAGACACTTAAGCCCGAGAACCTGCTTTGGATGAGAAAAACAATCCG 540  
598 CATTCGACATTTGGCATGCGCTCCCTGTCAGAGTGGGGAGAACCTCTCGAGACACCTG 657  
541 CATTCGACATTTGGCATGCGCTCCCTGTCAGAGTGGGGAGAACCTCTCGAGACACCTG 600  
658 CCGGTCCTCCCATTAATGCGGTCCAGAGTGAATTAAGGGGAAAAATGATGGCCGCG 717  
601 CCGGTCCTCCCATTAATGCGGTCCAGAGTGAATTAAGGGGAAAAATGATGGCCGCG 660  
718 GCGACATGTGGAGTGTGAGTCACTCTTCGCCCTCTGCGGGGCTCTGCCCTT 777  
661 GCGACATGTGGAGTGTGAGTCACTCTTCGCCCTCTGCGGGGCTCTGCCCTT 720  
778 TGATGACGACACCTCTCGGACGCTCTGAGAGAGTAAACGGGGCTCTTCCACATGCC 837  
721 TGATGACGACACCTCTCGGACGCTCTGAGAGAGTAAACGGGGCTCTTCCACATGCC 780  
838 CCACTTCATCTCCAGATGTCAGAGCTCTCGAGGGGAAATGTCAGAGTGAAGCCGA 897  
781 CCACTTCATCTCCAGATGTCAGAGCTCTCGAGGGGAAATGTCAGAGTGAAGCCGA 840  
898 AAAAGGCTCAGTCTGAGGCAATTCAGAAACATCTTGTGACTAGCGGGAA 951  
841 AAAAGGCTCAGTCTGAGGCAATTCAGAAACATCTTGTGACTAGCGGGAA 894

RESULT 7  
ABA08924  
ID ABA08924 standard; cDNA; 614 BP.

ABA08924;

11-JAN-2002 (first entry)

Human Ser/Thr kinase homologue-encoding cDNA, SEQ ID NO:700.

Human: cytokine; cell proliferation; tissue growth; immunomodulator; growth factor;  
inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;  
proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
chronic inflammatory condition; proliferative retinopathy;  
atherosclerosis; coronary heart disease; arterial ischaemia;  
bone disorder; osteoporosis; vascular growth disorder;  
tissue regeneration; wound healing; infection; immune disorder;  
cell culture; drug screening; gene therapy; anti-inflammatory;  
antiallergic; antiarthritic; haemostatic; antiarteriosclerotic;  
cytostatic; osteopathic; vasotropic; cardiatic; virucide; antibacterial;  
antifungal; vulnery; antilucer; ss.

Homo sapiens.

WO200157188-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US03800.

03-FEB-2000; 2000US-0496914.

27-APR-2000; 2000US-0560875.

(HSE-) HSEQ INC.

Tang YF, Liu C, Drmanac RT;

WPI; 2001-457740/49.

P-PSDB; AB11680.

PT Human proteins and DNA encoding sequences useful for preventing,  
PT treating or ameliorating a medical condition in a mammalian subject  
PT e.g. arthritis and cancer -  
XX  
PS Claim 1: Page 648; 1963pp: English.

XX Sequences AB10981-AB12330 represent 1350 novel human polypeptides, and  
CC sequences AB108225-AB109574 represent nucleic acids encoding them. The  
CC invention also relates to vectors and recombinant host cells comprising a  
CC nucleotide of the invention, methods of producing the novel polypeptides,  
CC antibodies against the polypeptides, methods of detecting the nucleotides  
CC or polypeptides in a sample, and methods of identifying compounds which  
CC bind to polypeptides of the invention. Although novel, many of the  
CC polypeptides of the invention have homology to known proteins, thereby  
CC giving an insight into their probable biological activities, and hence  
CC potential therapeutic applications. The polypeptides of the invention may  
CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC haematopoietic regulatory activity; tissue growth activity;  
CC immunomodulatory activity; activin- or inhibin-related activities;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis, cancer cell proliferation or metastasis.  
CC Depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical  
CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a cDNA encoding a  
CC novel human polypeptide of the invention.

XX Sequence 614 BP: 128 A; 226 C; 161 G; 99 T; 0 other;

Query Match 20.6%; Score 614; DB 22; Length 614;

Best Local Similarity 100.0%; Pred. No. 6,7e-111;  
Matches 614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1803 TCCCTACCGGTGAGAGATGTCAGAGTTCCTCCCGAGAGTGGCAAAAC 1862  
DB 1 TCCCTACCGGTGAGAGATGTCAGAGTTCCTCCCGAGAGTGGCAAAAC 60  
QY 1863 GCTCCGCTGCGGAGACTTCATCTCTCTGACAAAGAAATATTCCTCGTCTAA 1922  
DB 61 GCTCCGCTGCGGAGACTTCATCTCTCTGACAAAGAAATATTCCTCGTCTAA 120  
QY 1923 AGCAAAACCTCTAGCAGCATCAAGAGACATGTCATCTCTCTGATCCCA 1982  
DB 121 AGCAAAACCTCTAGCAGCATCAAGAGACATGTCATCTCTCTGATCCCA 180  
QY 1983 GCTGAGTACAGTGTGTGTACAGACCAAGTTCAGGCGGAGTACAAAGCCAGTGGCG 2042  
DB 181 GCTGAGTACAGTGTGTGTGTACAGACCAAGTTCAGGCGGAGTACAAAGCCAGTGGCG 240  
QY 2043 GCGCCCTCGTTCCTCAAAAGCCGCGCTTCAGGTGAGTCAAGTCTCTCTGAGGTC 2102  
DB 241 GCGCCCTCGTTCCTCAAAAGCCGCGCTTCAGGTGAGTCAAGTCTCTCTGAGGTC 300  
QY 2103 CAGAGCCCTCCCGGAGCGGAGCGAGGAGTGTGGATCTACTCTCCGTCACCTTCA 2162  
DB 301 CAGAGCCCTCCCGGAGCGGAGCGAGGAGTGTGGATCTACTCTCCGTCACCTTCA 360

QY 2163 CTCTCATCTCGGTTCCCGAGCCGTCGTTCAAGGAGTGTGGAGACCATCCAGGCACAGC 2222  
DB 361 CTCTCATCTCGGTTCCCGAGCCGTCGTTCAAGGAGTGTGGAGACCATCCAGGCACAGC 420  
QY 2223 TCCGTGAGCACTCATGAGACACCCCTCCGTCAGAGCCCTGGGAGAGAGAAAGGGGGCC 2282  
DB 421 TCCGTGAGCACTCATGAGACACCCCTCCGTCAGAGCCCTGGGAGAGAGAAAGGGGGCC 480  
QY 2283 AGACCCGGGCTGT 2342  
DB 481 AGACCCGGGCTGT 540  
QY 2343 CAGAGCTGAGCAGCTCTCCCGGAGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2402  
DB 541 CAGAGCTGAGCAGCTCTCCCGGAGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
QY 2403 AGGGAGCCCTCTG 2416  
DB 601 AGGGAGCCCTCTG 614

RESULT 8

ID ABK70216 standard; cDNA; 512 BP.

ABK70216;

15-JUL-2002 (first entry)

Human lung cancer associated cDNA SMO ID 87.

Human: ss; gene; lung cancer; cytostatic; tumour; vaccine.

Homo sapiens.

MO200224057-A2.

28-MAR-2002.

20-SEP-2001; 2001WO-US42232.

22-SEP-2000; 2000US-234837P.

10-OCT-2000; 2000US-229440P.

29-JUN-2001; 2001US-301928P.

(CORI-) CORIXA CORP.

Benson DR, Mohamath R, Lodes MJ;

WPI; 2002-372001/40.

New tumour lung proteins and nucleic acids encoding the proteins, useful  
PT as vaccines and for treating, preventing, diagnosing or monitoring lung  
PT cancer

Claim 1: Page 132-133; 189pp: English.

XX The invention relates to an isolated polynucleotide comprising a sequence  
CC selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),  
CC or their fragments, homologues, variants or complements and their encoded  
CC polypeptides. Also included are an expression vector comprising the  
CC polynucleotide operably linked to an expression control sequence; a host  
CC cell transformed or transfected with an expression vector of; an isolated  
CC antibody, or its antigen-binding fragment that specifically binds to the  
CC polypeptide; a method for detecting the presence of a cancer in a  
CC patient; a fusion protein comprising at least the polypeptide; an  
CC oligonucleotide that hybridises to the polynucleotide under moderately  
CC stringent conditions; a method for stimulating and/or expanding T cells  
CC specific for a tumour protein; an isolated T cell population comprising T  
CC cells prepared from the method of above; a composition comprising a first  
CC component consisting of carriers and immunostimulants, and a second  
CC component selected from the polynucleotides, proteins, antibodies, fusion



```

Db      509 TGTGGAGACCTCTCCGCGAGCCACACTACGCGTGCAGAGCTCATACGCGCGAGA 568
Oy      702 AATATATGCGCCGCGCGAGACATGTGAGTGTGAGTCTCTCTGCGTGC 761
Db      569 AGTACATGCGCGCGAGCGCGATGTCTGCTGCTGCGGATCTCTATGCGCTCC 628
Oy      762 TGGGGGCTCTGCGCTTGTATGACGACAACTCCGCCAGCTGCTGAGAGTGAACGG 821
Db      629 TGGGTCGCTGCGCTTGTGACGACAACTTGGCGCCAGCTGCTGAGAAAGTCAAGCGG 688
Oy      822 GCGTCTTCCAGACGCGCCCTTCTATTCCTCCATATTCGCGAGAGCTCTGAGGGAATGA 881
Db      689 GCGTCTTCCAGACGCGCTTGTGCGCGAGCTCCAGAGTCTGCTGCGCGCATGA 748
Oy      882 TCGAATGAGCCCGGAGAAAGAGCTAGTCTGAGCAATTCAGAAACATCTTGTACC 941
Db      749 TTGAGTGTCAATCCGAGACCGCGGCTCAGCGTGTGTAATCAACCGGCATCCGTGGCA 808
Oy      942 TAGCGGAGAAACGAGACGACCCGCTGCTGAGCAGCCCGCGCGGTTAGCCA 1001
Db      809 CAGCTGCGGCAAAAGG---GAGCTGAGCTGAGAGTGCATATGAGAGGTGGCAGA 865
Oy      1002 TGGCGAGCCTGCGCTCAGAGGAGAGTGCAGCCCGGCTCTAGAGAGCATGGCATAC 1061
Db      866 CACACGTTATTCACAGCCAGCCGCGGTGATCCGATGTGTTAACCGCATTTGCTGCG 925
Oy      1062 TGGGCTGCTTACGAGACCGCGAGAGCTGCATCGAGCTCGCATGAGAGAGAGAAC 1121
Db      926 TGGGCTGCTTACGAGAGAGAGAGAACTCATCCAGAACTGCTCATGTTCAAGTCAAA 985
Oy      1122 AAGAAAGATGATATATATCTCTTTTGTGATCGAAGAGAGCGGTATCCAGCTGTGAG 1181
Db      986 CGGAGAGAGGTATATATTTCTCTGCTGAGAGCGCAAAAGAGAGAGCTGCTGAGAG 1045
Oy      1182 ACCAGAGCTGCTCTCCCGGAA 1203
Db      1046 ATGATGATGAGATTGCGCAAAA 1067

RESULT 10
AAK94923
ID      AAK94923 standard; cDNA; 2954 BP.
XX      AAK94923;
AC      XX
XX      06-NOV-2001 (first entry)
DT      XX
XX      Human full-length cDNA, SEQ ID NO: 4157.
DE      XX
XX      Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
KW      XX
XX      Homo sapiens.
OS      XX
XX      EP1130094-A2.
PN      XX
XX      05-SEP-2001.
PD      XX
XX      07-JUL-2000; 2000EP-0114089.
PE      XX
XX      08-JUL-1999; 99JP-0194486.
PR      11-JAN-2000; 2000JP-0118774.
PR      02-MAY-2000; 2000JP-0183765.
XX      XX
XX      (HELI-) HELIX RES INST.
PA      XX
XX      Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI      Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX      WPI: 2001-524255/58.
DR      P-PSDB; AAM03956.
XX      PT
XX      830 Primers useful for synthesizing full length cDNA clones and their

```

```

PT      use in genetic manipulation -
XX      Claim 8; SEQ ID NO 4157; 1380bp + sequence listing; English.
PS      The invention relates to primers for synthesizing full length cDNA
XX      clones. 830 cDNA molecules encoding a human protein have been
CC      isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC      molecules have been determined. Primers for synthesizing the full length
CC      cDNA are useful for clarifying the function of the protein encoded by
CC      the cDNA. The full length clones were obtained by construction of full
CC      length enriched cDNA libraries that were synthesised by the oligo-capping
CC      method. The primers enable the production of the full length cDNA easily
CC      without any special methods. The present sequence is a full length
CC      human cDNA of the invention.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in CD-ROM format directly from EPO.
XX      XX
SQ      Sequence 2954 BP; 620 A; 970 C; 868 G; 496 T; 0 other;

Query Match      9.0%; Score 267.4; DB 22; Length 2954;
Best Local Similarity 55.6%; Pred. No. 7.6e-43;
Matches 514; Conservative 0; Mismatches 411; Indels 0; Gaps 0;

Oy      13 GTCCGCGACGCTGGCGGACCGCTGCGGAGCGCTGGGGGGGACCGGTGCGCGCG 72
Db      39 GTCTTCGCGGACGCTGCTGCGCGGCGCAAGATCGGAACCTCGACACGATGCGACCTT 98
Oy      73 GACCAAGGACACATGTCTGCGCGGCGCAAGAGAGTGGGGGCTCTCCGCGCTACA 132
Db      99 GGGCAATGCGCGCTCTCTCGGACAAAGGCCCGTCTGTCCAGCGCTCATCGGTGCGCG 158
Oy      133 CCTCCCGCACCCGCCACCCACCCACCCAGCACGCCCAATATGTGGGCCCTATCGGCT 192
Db      159 TTGCGGAACTCATGCGCTCTGCTGCGGAGAGACACCCACGCTGGGCAACTCGCGCT 218
Oy      193 GGAGAAAGCGCTGGCAAGAGACAGAGGCTGTAACTCGGGGTCCATCGATCAC 252
Db      219 GCTGAGGACATTGGGAAGGCAACTTGGCAAGTCAAGCTGCTGCGGACATCTCAC 278
Oy      253 GGGTCAGAAAGTCCCATCAAGATCTGTAACCGGAGAGAGCTGCGAGTGGTGTGAT 312
Db      279 TGGTCGGAGAGTTCGCATCAAGATTCAGCAAAACCCAGCTGTAATCCACAGCTGCA 338.
Oy      313 GAAGTGGAGCGGAGATGCCATCTGAAGCTCATGAAACCCACCATGCTCTCAAGCT 372
Db      339 GAAGCTGTTCCGAGAAAGTCCGATCATGAAGGCTTAACCCACCATGCTGAAGCT 398
Oy      373 CCACGAGCTCTAGAGAAAGAAATTTGTACCTGTTCTGAGAGACGCTGCGGGGG 432
Db      399 CTTTGAAGTGTATGAGACTGAGAAAGACGCTGTACTGTGATGAGATGCGAAGTCTGG 458
Oy      433 TGAGCTATTGACTACTGTGTAAGAGAGGAGAGTACGCGCCAGAGAGGCGCGAAAGTT 492
Db      459 AGAAGTGTGACTACTGCTGTGTCGATGCGCATGAGGAGAGAGAGAGTCCAGGCCAA 518
Oy      493 CTTCCGCGAGATTGTGTCTGCTGAGACTTCTGCCACAGCTACTCTCCACAGAGA 552
Db      519 GTTCCAGACAGATTGTTGCGCTGTGACATTTCTACAGAAATATTTGTACACAGGA 578
Oy      553 CTTAAAGCCCGAGAACTGCTTTGTGATGAGAAAGAAACATCGCGCATTTGAGACTGGG 612
Db      579 CCGAAGGCTGAGAAACCTCTTGTGATGCGGAGCCCAACATCAAGATTGCTGACTTTGG 638
Oy      613 CATGGGCTCCCTCAGGTGGGAGACGCTCTCTGAGAGACGATCGGGGTCCGCCATTA 672
Db      639 CTTACAGAAAGATTCACGCTGGGATGGAAGCTGGACACGTTCTGCGGAGACCCCATTA 698
Oy      673 TGGCTGTCAAGAGTGTATTAAGGGGAAATATATGATGCGCCGCGGACATGTGAG 732
Db      699 TGGCGGCGCGAGACTGTTTACAGCAAGAGTACAGAGGCGGAGGATGACTGTGAG 758
Oy      733 CTGTGAGTCACTCTCTTCCCGCTGCGTGGGGGCTGCTGCTTGTATGAGCAACCT 792

```



DB 679 TCCGCCGCCGAGCTGTTTACAGGCAAGAACTACAGACGCGCCGAGGCTGACATCTGGAG 738  
 QY 733 CTGTGAGTCACTCTCTGCGCCCTGCTGTTGGGGGCTCTGCCCTTGTATGACGACACCT 792  
 DB 739 CTTGGGAGTCACTCTGTTACACCCCTGTCAGCGGCTCCCTGCCCTTTCGACGGGACACACT 798  
 QY 793 CCGCCAGCTCTGAGAAAGGTGAACGGGGGCTCTTCACATGCCCTTCATTCCTCC 852  
 DB 799 CAGGAGCTCTGGAGGAGTACTACAGAGGAAAGTACCGGGTCCCTTTCATCATGTCAAC 858  
 QY 853 AGATTGCGACAGGCTCTGAGGAGATGATGAGTGGAGCCGAAAAGGCTGAGCT 912  
 DB 859 AGACTGTGAGAGCACTCTGCGGAGATTTTGTGCTGACACCACTTAACGCTGACTCT 918  
 QY 913 GGAGCAATTCAGAAACATCTTGG 937  
 DB 919 CGAGCAATCATGAAAGCAAAATGG 943

RESULT 12  
 ABA05739  
 ID ABA05739 standard; cDNA; 3312 BP.  
 AC ABA05739;  
 DT 04-MAR-2002 (first entry)  
 DE Human neuronal serine threonine protein kinase 9B5 coding sequence.  
 XX Human: neuronal serine threonine protein kinase; apoplexy; cytosolic;  
 KM cerebroprotective; neuroprotective; antiparkinsonian; immunosuppressive;  
 KM viral; anti-HIV; antirheumatic; antineoplastic; antitumor; cancer;  
 KM cerebral infarction; neurological disease; immunological disease; atopy;  
 KM infection; leukaemia; polyarthritis; Crohn's disease; ulcerative colitis;  
 h9B5; ss.  
 KM  
 XX Homo sapiens.  
 OS  
 FH Location/Qualifiers  
 FT CDS 64..3280  
 FT /\*tag- a  
 FT /product- "h9B5"  
 FT 3275..3280  
 FT /\*tag- b  
 FT  
 FT polyA\_signal  
 FT /tag- b  
 PN  
 XX W0200188108-A1.  
 PD 22-NOV-2001.  
 PF 17-MAY-2001; 2001MO-EP05660.  
 PR 17-MAY-2000; 2000DE-1024171.  
 PR (BADI) BASF-LYNX BIOSCIENCE AG.  
 PA  
 XX Schneider A, Klausner B, Fischer A, Newrzella D, Goetz B;  
 PI Rosner M, Eisenhardt G, Kuner R, Trutzel A, Kammandel B;  
 PI Jomana Naim S, Schwaninger M;  
 XX MPI; 2002-055696/07.  
 DR P-PSDB; ABB04433.  
 DR  
 XX New polynucleotide, useful for the diagnosis, treatment and prevention  
 PT of e.g. apoplexy, tumors and autoimmune disease, comprises a  
 PT polynucleotide encoding the neuronal protein kinase -  
 PS Claim 1; Page 54-55; 75pp; German.  
 XX The present invention provides the protein and coding sequences of two  
 CC neuronal serine threonine protein kinases each from the mouse and human.  
 CC These are designated 9B5 and 9B5.b. Detection of their expression levels  
 CC can be used to diagnose the risk of apoplexy and in prognosis of cerebral

CC infarction. They can also be used to identify agents useful in the  
 CC prevention, treatment and diagnosis of neurological diseases, including  
 CC stroke, multiple sclerosis, Parkinson's disease, amyotrophic lateral  
 CC sclerosis, tumors, especially carcinoma, immunological disorders,  
 CC including autoimmune diseases, atopy, viral (including human immune  
 CC deficiency virus) infections, acute or chronic leukocytic or myeloid  
 CC leukaemia, primary chronic polyarthritis, Crohn's disease and ulcerative  
 CC colitis. The present sequence is the human 9B5 coding sequence.  
 XX  
 SQ Sequence 3312 BP; 710 A; 1071 C; 954 G; 577 T; 0 other;  
 Query Match 9.0%; Score 267.4; DB 24; Length 3312;  
 Best Local Similarity 55.6%; Pred. No. 7.9e-43;  
 Matches 514; Conservative 0; Mismatches 411; Indels 0; Gaps 0;  
 QY 13 GTCCGGGAGACGCTGGGCGAGCGCTGGGCGAGCGCTGGGGGGGACCGGCTGGCGCG 72  
 DB 66 GTCTTCGCGGACCGGTGCTGTGGCCCGGCGCAACGATCGGAACTCGACACGATGGACCTT 125  
 QY 73 GACCAAGGACCAATGTCGTCCGGGCGCAAGAGGAGGTGGGGCTCTCCGCTTACCA 132  
 DB 126 GGCGAGTGGCGCTCTCTGGAACAAGGCGCTCTGTCAGCGCTCATGAGTGGCGCG 185  
 QY 133 CTTCCCGGACCCCGGACCCCGGACCCCGGACCCCGGACCCCGGACCCCGGACCCGCT 192  
 DB 186 TTCCCGGACCTGATCGCTCTGTCGTCGCGGAGGACCGCCGACGCTGGCACTACCGCT 245  
 QY 193 GGAGAAAGCGCTGGGCGCAAGAGACAGAGGCTGTTAACTGGGGGCGGACGCTATAC 252  
 DB 246 GCTGAGAGACCTATGGGAGAGGCACTTTGCCAAAGTCAAGCTGGCTGGCACATCTTAC 305  
 QY 253 GGGTCAGAGAGCGCGGCAATCAAGATGTAACCGGGAGAGCTGTGGAGTGGCTGAT 312  
 DB 306 TGGTCGGGAGTGGCTCAATCAAGATTAATGACAAACCGGAGTGAATCCGACAGCTCTA 365  
 QY 313 GAAGGTGAGAGGGGAGATTCGCACTGGAAGCTATGCAACCCGACATGTCCTAACCT 372  
 DB 366 GAAGCTGTTCGAGAACTCCGATCATGAAAGGCTTAAACCCGACCAATGCTGAACCT 425  
 QY 373 CCAGAGCTCTAGAGAAACAATAATTTTACCTGTTCTGGAGACGCTCTGGGGGG 432  
 DB 426 CTTTGAAGTATTTAGAGCTAGAGAGAGAGCGCTTACTGTTATGAGTACGCAAGTCTGG 485  
 QY 433 TGAGCTATTCGACTACCTGTGTAAGAAAGGGAGAGTACGCCGCAAGAGGCGCGAACTT 492  
 DB 486 AGAAGTGTGACTACCTGCTGTCGATGCGGCGGATGAAGAGAGAGAGTCTGAGGCA 545  
 QY 493 CTTCCCGGACATTTGCTGCTGGCTGAGCTTCTGCAAGCTACTTCATGCTGCAAGAGA 552  
 DB 546 GTTCCGACAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 605  
 QY 553 CTTAAAGCCCGAGAACTGCTTTGATGAGAAACCAATCCGATTTGAGACTTGG 612  
 DB 606 CTTGAAGCTGAGAGACTCTTCTGATGATCCGAGGCGCAACATCAAGATTTGCTGACTTTGG 665  
 QY 613 CATGGCTCCCTGAGAGTGGGGGACAGCCCTCTGAGACACAGCTGGGGTCCCGCCATTA 672  
 DB 666 CTTGAGCAACGAGTTCACGCTGGGATGAGAGCTGAGACGTTTGGGGGAGGAGGCCCATTA 725  
 QY 673 TGGCTGTCCAGAGGTGATTAAGGGGGAATAATATGATGAGCCCGGCGGACGATGTGAG 732  
 DB 726 TGGCGCCCGGAGGTGTTTACAGGCAAGAAATAGAGGCGGCGGAGGTGAGACTCTGAG 785  
 QY 733 CTGTGAGTCACTCTCTGCGCCCTGCTGTTGGGGGCTCTGCTTTGTATGACACACCT 792  
 DB 786 CTTGGGAGTCACTCTGATACACCTCTGTCAGGCGCTCCCTGCTTTCGACGGGACACACT 845  
 QY 793 CCGCCAGCTCTGAGAGAGGTGAACCGGGGCTTTCACATGCCCTTCATTCCTCC 852  
 DB 846 CAGGAGCTCGGAGAGGAGTACTCAGAGGAGAGTACCGGGGCTCTTCTCATGATCAAC 905  
 QY 853 AGATTGCGAAGCTCTGAGGAGGAAATGATGAAATGAGCCCGGAAAAGGCTAGCT 912

DB 906 AGACTGTGAGAGCATCTCCGCGAGATTGTCCTGACACCCAGCTAAACGCTGACTCT 965  
 QY 913 GGAGCAAAATTCAGAAACATCTGG 937  
 DB 966 CGAGCAAAATTCAGAAACATG 990

RESULT 13  
 ABA05740  
 ID ABA05740 standard; cDNA; 3392 BP.

AC ABA05740;

DT 04-MAR-2002 (first entry)

DE Human neuronal serine threonine protein kinase 9B5\_b coding sequence.

KW Human: neuronal serine threonine protein kinase; apoptosis; cytostatic;  
 KW cerebroprotective; neuroprotective; antiparkinsonian; immunosuppressive;  
 KW virucide; anti-HIV; antiarthritic; antiinflammatory; anticancer; cancer;  
 KW cerebral infarction; neurological disease; immunological disease; atopy;  
 KW infection; leukaemia; polyarthritides; Crohn's disease; ulcerative colitis;  
 KW h9B5\_b; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 64..2130

FT /tag- a

FT /product- "h9B5"

FT polyA\_signal 3355..3360

FT /tag- b

PN WO200188108-A1.

PD 22-NOV-2001.

PF 17-MAY-2001; 2001WO-EP05660.

PR 17-MAY-2000; 2000DE-1024171.

PA (BADI) BASF-LYNX BIOSCIENCE AG.

PI Schneider A, Klausner B, Fischer A, Newrzella D, Goetz B;

PI Rosner M, Eisenhardt G, Kumer R, Trutzel A, Kammandel B;

PI Jomana Naim S, Schwabinger M;

DR MPI: 2002-055696/07.

DR P-PSDB; ABB04434.

PS New polynucleotide, useful for the diagnosis, treatment and prevention

PT of e.g. apoptosis, tumors and autoimmune disease, comprises a

PT polynucleotide encoding the neuronal protein kinase -

XX Claim 1; Page 55-56; 75pp; German.

XX The present invention provides the protein and coding sequences of two

CC neuronal serine threonine protein kinases each from the mouse and human.

CC These are designated 9B5 and 9B5\_b. Detection of their expression levels

CC can be used to diagnose the risk of apoptosis and in prognosis of cerebral

CC infection. They can also be used to identify agents useful in the

CC prevention, treatment and diagnosis of neurological diseases, including

CC stroke, multiple sclerosis, Parkinson's disease, amyotrophic lateral

CC sclerosis, tumours, especially carcinoma, immunological disorders,

CC including autoimmune diseases, atopy, viral (including human immune

CC deficiency virus) infections, acute or chronic leukocytic or myeloid

CC leukaemia, primary chronic polyarthritides, Crohn's disease and ulcerative

CC colitis. The present sequence is the human 9B5\_b coding sequence.

CC Sequence 3392 BP; 726 A; 1099 C; 972 G; 595 T; 0 other;

XX Query Match

XX Best Local Similarity 9.0%; Score 267.4; DB 24; Length 3392;

XX 55.6%; Pred. No. 7.9e-43;

Matches 514; Conservative 0; Mismatches 411; Indels 0; Gaps 0;

QY 13 GTCCGCGAGCGCTGGGCGGACGCGTGCGGCGGAGCGCTGGGGGGGACCGGTGGCGG 72  
 DB 66 GTCTTGGCGGAGCGTCTGCGCGCGCGGCGCAACATCGAAGCTGGACACGCGTACCTT 125  
 QY 73 GACCAAGGCGACCATCTCTCGCGGGGCGCAAGAGGAGGTGGGGGCTCCCGCTACCA 132  
 DB 126 GGGCAGTGGCGCGCTCTCGGCAAAAGGCCCGCTCTGTGTCACAGCGCTACCTGGGTGCGG 185  
 QY 133 CCGCCCCCGACCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCT 192  
 DB 186 TTGCGGAACTCCATCGCT 245  
 QY 193 GGAGAAAGCGCTGGGCGCAAG 252  
 DB 246 GCGAG 305  
 QY 253 GGGTCAGAAAGTCCGCTCAAGATCGTGAACCGGAGAGAGAGAGAGAGAGAGAGAGAG 312  
 DB 306 TGGTCGGAGAGTTCGATCAAGATTCGCAAAACCCAGCTGAATCCAGAGAGCTGCA 365  
 QY 313 GAAGTGGAGCGGAGAGATCGCCATCTGAAAGCTCATGAAACCCAGATGCTCCCAAGCT 372  
 DB 366 GAAGCTGTTCCGAGAAAGTCCGATCATGAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 425  
 QY 373 CCACGAGCTCTACGAG 432  
 DB 426 CTTTGAAGTGTATGAGACAG 485  
 QY 433 TGAAGTATTCGATACCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492  
 DB 486 AGAAGTGTGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545  
 QY 493 CTTCCGCGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552  
 DB 546 GTTCCGAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 605  
 QY 553 CCTAAAGCCGAG 612  
 DB 606 CCGTGAAGGCTGAGAAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 665  
 QY 613 CATGGCGCTCCGAG 672  
 DB 666 CTTACAGCAAG 725  
 QY 673 TGGGTGTCCAGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 732  
 DB 726 TGCCTGCGCGGAG 785  
 QY 733 CTGTGAGTCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792  
 DB 786 CTTGGAGTCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 845  
 QY 793 CCGCAGCTGTGAG 852  
 DB 846 CAAGAGCTGCGGAG 905  
 QY 853 AATGTTCAG 912  
 DB 906 AGACTGTGAG 965  
 QY 913 GGAGCAAAATTCAGAAACATCTGG 937  
 DB 966 CGAGCAAAATTCAGAAACATG 990

RESULT 14

AAF44655  
 ID AAF44655 standard; cDNA; 1594 BP.

XX AAF44655;

XX



XX Schneider A, Klausner B, Fischer A, Newrzella D, Goetz B;  
 PI Rosser M, Eisenhardt G, Kuner R, Trutzel A, Kammandel B;  
 PI Jomana Naim S, Schwaninger M;  
 XX  
 DR MPI: 2002-055696/07.  
 P-PSDB: ABB04431.

XX New polynucleotide, useful for the diagnosis, treatment and prevention  
 of e.g. apoptosis, tumors and autoimmune disease, comprises a  
 polynucleotide encoding the neuronal protein kinase -

PS Claim 1: Page 52-53; 75pp: German.

CC The present invention provides the protein and coding sequences of two  
 CC neuronal serine threonine protein kinases each from the mouse and human.  
 CC These are designated 985 and 985.B. Detection of their expression levels  
 CC can be used to diagnose the risk of apoptosis and in prognosis of cerebral  
 CC infarction. They can also be used to identify agents useful in the  
 CC prevention, treatment and diagnosis of neurological diseases, including  
 CC stroke, multiple sclerosis, Parkinson's disease, amyotrophic lateral  
 CC sclerosis, tumours, especially carcinoma, immunological disorders,  
 CC including autoimmune diseases, atopy, viral (including human immune  
 CC deficiency virus) infections, acute or chronic leukocytic or myeloid  
 CC leukaemia, primary chronic polyarthritis, Crohn's disease and ulcerative  
 CC colitis. The present sequence is the murine 985 coding sequence.

XX Sequence 3170 BP; 714 A; 988 C; 897 G; 571 T; 0 other;

Query Match 8.7%; Score 259.4; DB 24; Length 3170;

Best Local Similarity 57.4%; Pred. No. 2.9e-41; Mismatches 346; Indels 0; Gaps 0;

QY 174 ATGTGGCCCTATCGCTGAGAGAGCGCTGGCCAAAGACAGACAGCGCTGTTAAAC 233  
 DB 80 ATGTGGCCCACTATAGCTGTGTAGAACCATCGGGAAGGCACTTGCCAAAGTCAAGC 139  
 QY 234 TCGGGGTCCTCATGATCAGGGGTGAGAGTCCCATCAAGATGCTGAACCGGGAAGC 293  
 DB 140 TGGCTGGGCTATTCCTCAGCGGCGGGAGGCTGCTATTAAGTATGATTAAGCCAGC 199  
 QY 294 TGTGGAGTGGTGTGATGAGAGGTGAGGAGGAGATGCGCATCTGAAAGCTATGAGAC 353  
 DB 200 TGAACCCCACTAGCTTCAAGAGCTGTTCAGAGAGATCCGATTAAGAGGACTCAAC 259  
 QY 354 ACCCAGATGCTCTCAAGCTCCAGACGCTTAAGAGAACAAATATTTGTAACCTGTTTC 413  
 DB 260 ACCCAACATCGTGAAGCTTTTGAAGTGAAGAGAGGAGAGAGCGTATACCTGGTGA 319  
 QY 414 TGAAGACGCTCTGGGGGTGAGCTATTCGACTACCTGGTAAAGAGGGGAGACTGACGC 473  
 DB 320 TGAATACGCTAGCGGAGGAGAACTGTTGACTCTGCTGCGACGCGCATGAAGG 379  
 QY 474 CCAAGAGGCGCCGAATGTTCTCCGAGATGTTGCTGCGGTGAGCTTGCACAGCT 533  
 DB 380 AGAAGAGGCTCAAGCAAGTGTGGGCAAGATCTGTCAAGCCGTGACTACTGTCTCAGA 439  
 QY 534 ACTCATCTGCGCACAGAGACTAAAGCCGAGAACTGCTTTTGGATGAGAAAAACAACA 593  
 DB 440 AGAACATTTGACAGAGGATCTAAAGCTGAAAACCTGTTGCTGATGCCAGGCCAACA 499  
 QY 594 TCGCATTTGACACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 653  
 DB 500 TCAAAATCGCGACTTCGGCTTCAGCAATGAGTTCACGCTGCGCTCCCAAGCTGAGACCT 559  
 QY 654 GCTGGGGTCCCGCATTTATGCGTCCAGAGGATTAAGGGGAAAAATATATGATGACC 713  
 DB 560 TCTGTGGAGACCCCATATACCGCGCCAGAGCTGTTCAGAGGCAAGATATGATGAGGC 619  
 QY 714 GCGGGGACAGATGTGAGCTGTGAGTCACTCTTCCCTGCTGCTGCTGCTGCTGCTGCT 773  
 DB 620 CAGAGGTGAGACATCTGAGGCTGGGTCTCATCTGTAACGCTGCTGCTGCTGCTGCTGCT 679

QY 774 CTTTGTAGACGACAACTCTCCGCACTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 833  
 DB 680 CTTTCATGGGACAACTCTCAAGAGCTGGGAGCGCAATCTCAGAGAGGAGGAGGAGGAG 739  
 QY 834 TGGCCCATCTCATCTCCGAGATTGGCAGAGCCCTCGAGGGGAGATGATCAACTGGAGC 893  
 DB 740 TCCCTTCTACATGTCTACAGACTGCGAGAGCATTCGCGGATTTCTGTGCTGTAAC 799  
 QY 894 CCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTTGTACTAGCGGGAAC 953  
 DB 800 CCGCAAAAGCTGTACTCTGAGCAAAATTCAGAAACATTCAGAAACATTCAGCAACTGCTATG 859  
 QY 954 AGAGCCAGACCCGCTGCTGAGCCAGCCCTG 986  
 DB 860 AGGCTGAGAGCTGAAGCCATACACGAGGCTG 892

Search completed: April 21, 2003, 20:31:16  
 Job time : 428.371 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 23:05:46 ; Search time 182.234 Seconds

(without alignments)  
16425.361 Million cell updates/sec

Title: US-10-003-690-1

Perfect score: 2980  
Sequence: 1 gtcacccacgctccgcg.....aaaaaaaaagggcgccgcta 2980

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	965	32.4	2007	12 US-10-054-579-1	Sequence 1, Appl1
2	961	32.2	2025	9 US-09-842-582-3	Sequence 3, Appl1
3	961	32.2	2219	9 US-09-842-582-1	Sequence 1, Appl1
4	851.4	28.6	1827	12 US-10-054-579-3	Sequence 3, Appl1
5	512	17.2	512	10 US-09-960-253-87	Sequence 87, Appl1
6	296.8	10.0	2173	9 US-10-102-558-1	Sequence 1, Appl1
7	265.8	8.9	3270	10 US-09-835-081-1	Sequence 1, Appl1
8	257.8	8.7	316	9 US-09-764-868-191	Sequence 191, App
9	240.6	8.1	2112	10 US-09-919-585-4	Sequence 4, Appl1
10	240.6	8.1	2222	10 US-09-919-585-7	Sequence 7, Appl1
11	225.4	7.6	1539	9 US-09-938-842A-2657	Sequence 2657, Ap
12	221.4	7.4	2352	10 US-09-815-915-3	Sequence 3, Appl1
13	221.4	7.4	2968	10 US-09-815-915-1	Sequence 1, Appl1
14	215	7.2	275	10 US-09-960-352-5086	Sequence 5086, Ap
15	207	6.9	5963	10 US-09-799-875-3	Sequence 4, Appl1
16	202.8	6.8	3609	10 US-09-799-875-6	Sequence 6, Appl1
17	196.8	6.6	1539	9 US-09-938-842A-162	Sequence 162, App
18	196	6.6	3154	10 US-09-919-585-19	Sequence 19, Appl1
19	181	6.1	3079	10 US-09-834-975-974	Sequence 974, App

20	179.4	6.0	2271	10 US-09-919-585-1	Sequence 1, Appl1
21	178	6.0	2706	10 US-09-919-585-10	Sequence 10, Appl1
22	177.6	6.0	1902	10 US-09-801-368-337	Sequence 337, App
23	174	5.8	1884	10 US-09-963-159-3	Sequence 3, Appl1
24	174	5.8	3353	10 US-09-963-159-1	Sequence 1, Appl1
25	172.4	5.8	1266	9 US-09-938-842A-1288	Sequence 1288, Ap
26	167.2	5.6	4557	10 US-09-801-368-151	Sequence 151, App
27	157	5.3	267	10 US-09-983-965-2108	Sequence 2108, App
28	153.8	5.2	6828	10 US-09-780-949-1	Sequence 1, Appl1
29	153.4	5.1	2884	10 US-09-780-949-5	Sequence 5, Appl1
30	152.6	5.1	1986	10 US-09-780-949-3	Sequence 3, Appl1
31	147	4.9	2453	9 US-09-974-298-111	Sequence 111, App
32	147	4.9	2470	10 US-09-870-937-9	Sequence 9, Appl1
33	147	4.9	2470	10 US-09-954-456-1121	Sequence 1121, Ap
34	147	4.9	2470	10 US-09-880-107-1725	Sequence 1725, Ap
35	139.8	4.7	2162	10 US-09-822-849A-483	Sequence 483, App
36	137	4.6	1005	9 US-10-142-356-4	Sequence 4, Appl1
37	133	4.5	1326	9 US-09-938-842A-72	Sequence 72, Appl1
38	130.2	4.4	2519	9 US-10-071-766-82	Sequence 82, Appl1
39	126.8	4.3	1326	9 US-09-938-842A-2313	Sequence 2313, Ap
40	126.8	4.3	1635	10 US-09-880-107-2340	Sequence 2340, App
41	119.2	4.0	3594	9 US-09-836-392-7	Sequence 7, Appl1
42	118.6	4.0	610	9 US-10-040-739-475	Sequence 475, App
43	116.6	3.9	1750	9 US-10-142-356-1	Sequence 1, Appl1
44	115.8	3.9	3061	10 US-09-880-107-2146	Sequence 2146, App
45	112	3.8	1308	9 US-09-938-842A-1559	Sequence 1559, Ap

## ALIGNMENTS

```
RESULT 1
US-10-054-579-1
; Sequence 1, Application US/10054579
; Patent No. US20020137913A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; TITLE OF INVENTION: NO. US20020137913A1 Human Kinases and Polynucleotides Encodi
; FILE REFERENCE: LEX-0300-USA
; CURRENT APPLICATION NUMBER: US/10/054,579
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/263,378
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2007
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-054-579-1

Query Match      32.4%; Score 965; DB 12; Length 2007;
Best Local Similarity 70.0%; Pred. No. 2.1e-216;
Matches 1464; Conservative 0; Mismatches 440; Indels 195; Gaps 6;

QY 159 CCCAGCAGCCCAATATATGTCGTCATGCGGCGGCGGAGAGACGCTGGGCAAGAGACAGA 218
      |||||
DB 29 CCGAGCAGCCGCGAGTATGTTGGGCGCTACCGCGTGGAGAGACGCTGGGCGAGAGAG 88
      |||||

QY 219 CAGGCGTGTAACTCGGGGTCACCTGCATCAGCGGTCAGAGAGTGCCTCAAGATCG 278
      |||||
DB 89 CAGGCGTGTGAAGCTGGGGGTCACCTGCATCAGCGGTCAGAGAGTGCCTCAAGATCG 148
      |||||

QY 279 TGAACCGGAGAGAGCTGTGAGTGTGAGTGAAGAGTGGAGAGTGGCATCG 338
      |||||
DB 149 TCAACCGTGAAGAGCTGACGAGTGTGATGAAGAGTGGAGAGTGGCATCG 208
      |||||

QY 339 TGAAGCTCATGAGACACCACTGTCTCAAGCTTCACGAGAGTGTAGAGAGAGAAAT 398
      |||||
DB 209 TGAAGCTCATGAGACACCACTGTCTCAAGCTTCACGAGAGTGTAGAGAGAGAAAT 268
      |||||

QY 399 ATTGTACCTGCTTTCGAGACAGCTCTCGGGGGTGTAGCTATTGACTACTGCTGAAGA 458
      |||||
```

||||| 269 ATTTGATCGGTGCTGAGAACACGTGTGAGTGGTGTGAGCTTTGACTACCTGTGGAAGA 328  
||||| 459 AGGGAGAGATGAGCCCGGAGAGAGCCGAAAGTTCTTCCGCAATGTTGTGTGGCTGG 518  
||||| 329 AGGGAGAGCTGAGCCCTAGAGAGGCTGGAAATTTCTTCCGCAATATCTCTGTGGCTGG 388  
||||| 519 ACTTCTGCCACAGCTACTCATCTGTGACAGAGACCTTAAAGCCCGAGAACCTGTCTTTGG 578  
||||| 389 ACTTCTGCCACAGCTACTCATCTGTGACAGAGGATCTGAAACCTGAAACCTCTGTGG 448  
||||| 579 AAGAGAAACAAATCCGATTTGAGAGACCTTGGCATGGCTGCTCCATGAGTGGGGGACA 638  
||||| 449 AGGAGAGAAACAAATCCGATTTGAGAGACCTTGGCATGGCTGCTCCATGAGTGGGGGACA 508  
||||| 639 GCTCTCTGAGAGACAGCTGCGGGTTCCTCCATTTATGCTGTCCAGAGATTAAGGGGG 698  
||||| 509 GCTCTGAGAGACAGCTGCGGGTTCCTCCATTTATGCTGTCCAGAGATTAAGGGGG 568  
||||| 699 AAAAAATATGATGCGCCCGGAGACATGTGAGAGCTGTGAGATCATCTCTTTCGCTTGC 758  
||||| 569 AAAAAATATGATGCGCCCGGAGACATGTGAGAGCTGTGAGATCATCTCTTTCGCTTGC 628  
||||| 759 TGTGAGGAGCTGTGCTTGTATGAGAGACAACTCCGCGAGCTGTGAGAGATTAAGGGGG 818  
||||| 629 TGTGAGGAGCTGTGCTTGTATGAGAGACAACTCCGCGAGCTGTGAGAGATTAAGGGGG 688  
||||| 819 GGGGCTCTTCCATGAGCCCGGAGACATGTGAGAGCTGTGAGAGATTAAGGGGG 878  
||||| 689 GGGGCTCTTCCATGAGCCCGGAGACATGTGAGAGCTGTGAGAGATTAAGGGGG 748  
||||| 879 TGTATGAGTGTGAGAGCCCGGAGACATGTGAGAGCTGTGAGAGATTAAGGGGG 938  
||||| 749 TGTATGAGTGTGAGAGCCCGGAGACATGTGAGAGCTGTGAGAGATTAAGGGGG 808  
||||| 939 ACTTATGAGGAG 998  
||||| 809 ATTTATGAGGAG 862  
||||| 999 CCATGCGGAGAGCTGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1058  
||||| 863 AGATGCGGAG 922  
||||| 1059 CACTGCGGAG 1118  
||||| 923 CACTGCGGAG 982  
||||| 1119 ACCAAGAAAGATGAT 1178  
||||| 983 ACCAAGAAAGATGAT 1042  
||||| 1179 AGGAG 1238  
||||| 1043 AGGAG 1102  
||||| 1239 CCATGCGGAG 1298  
||||| 1103 CCATGCGGAG 1162  
||||| 1299 TCACGAGATGCGGAG 1358  
||||| 1163 TCACGAGATGCGGAG 1213  
||||| 1359 AGCAG 1418  
||||| 1214 AGCAG 1273  
||||| 1419 TAAGCAG 1478  
||||| 1274 TCACGAG 1286  
||||| 1479 CTCGAG 1538

1287 ----- 1286  
1539 GTGGGGGCGCGGAG 1598  
1287 ----- 1366  
1599 CCGCAGAGCTCCCGGAG 1658  
1327 ----- 1357  
1659 GGGCAGAGCTCCCGGAG 1718  
1358 AGGAG 1411  
1719 GGGAG 1778  
1412 GAGGGGTGCGCTGAG 1471  
1779 GCTTTCACCGGAG 1838  
1472 GCTTTCACCGGAG 1531  
1839 AGTCTCCCGGAG 1898  
1532 AGTCTCCCGGAG 1591  
1899 AAGAACAAATATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1958  
1592 AGGAG 1651  
1959 TCATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2018  
1652 TCATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711  
2019 GGGGAG 2078  
1712 GGGGAG 1771  
2079 TGAGATGAG 2138  
1772 TGAGATGAG 1813  
2139 GTGGATGAG 2198  
1814 AGGAGATGAG 1873  
2199 TGAG 2258  
1874 TGAG 1933  
2259 TGAG 2277  
1934 TGAG 1952

RESULT 2  
US-09-842-582-3  
Sequence 3, Application US/09842582  
Patent No. US20020155570A1  
GENERAL INFORMATION:  
APPLICANT: Millennium Pharmaceuticals, Inc.  
TITLE OF INVENTION: 2246, NOVEL PROTEIN KINASE MOLECULES AND  
TITLE OF INVENTION: USES THEREFOR  
FILE REFERENCE: 38155-20054.00  
CURRENT APPLICATION NUMBER: US/09/842,582  
PRIORITY FILING DATE: 2001-04-25  
PRIOR APPLICATION NUMBER: US 60/199,391  
PRIORITY FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 2025

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-842-582-3

```

Query Match	32.2%;	Score 961;	DB 9;	Length 2025;
Best Local Similarity	70.1%;	Pred. NO. 1.8e-215;		
Matches 1477;	Conservative	0;	Mismatches 435;	Indels 195; Gaps 6

QY	159	CCACACAGCCCAATATGTGGGCCCCCTATCGCTGGAGAAAGACGCTGGCCAAAGACAGA	218
Db	29	CGCAGCACGCGCAAGTATGTGTGGGCCCTACCGGCTGGAGAAAGACGCTGGCCAAAGGGCCAGA	88
QY	219	ÇAGGGCTGGTTAAACTCGGGGTCACATGCAATCACGGGGTGAAAGTCCCATCAAGATCG	278
QY	279	TGAACCGGGAGAACTGTGCGAGTGGGTGCTGATGAAAGTGGAGCGGGAGATGCCCATCC	338
Db	149	TCACACCGTAGAAGCTCACGCGAGTGGTCTGTATGAAAGGTGGAGCGGGAGATCCCGATCC	208
QY	339	TGAAGCTCATGCACACCCACATGTCTCAAGCTCCACGACGTCCTACGAGAACAAAGAAAT	398
Db	209	TGAAGCTCATGTAGACACCCCCACGTCCTTAAGCTGCACGACGTTTATGAAACCAAAAAAT	268
QY	399	ATTGTGTACCTGGTTCTGAGCACGTCGCGGGGTGAGCTATTGACTACCTGGTAAAGA	458
Db	269	ATTGTGTACCTGGTGTAGAACACGTTGTCAGTGGTGGTGGCTTCGACTACCTGGTGAAGA	328
QY	459	AGGGAGACTGACGCCCAAGGAGGCGCCGAAAGTCTTTCGCCCAAGATTGTGTCGCTGG	518
Db	329	AGGGAGGCTGACGGCTTAAGGAGGCTCGGAAAGTCTTCGCGGACAGATATCTCTGCGCTGG	388
QY	519	ACTTGTGCGACAGTACTCATCTGTGCGACAGAGACCTTAAGCCCGAGACACTGCTTTGG	578
Db	389	ACTTGTGCGACAGCCACTTCATATGCGACAGAGSATGTGAAACTGAAAACCTCCTGCTGG	448
QY	579	ATGAGAAAAACAACATCCGCAATTGACAGCTTGCGCATGGCGTCCCTGACAGTGGGGACA	638
Db	449	ACGAGAAACAACAATCCCATTCGACAGACTTGGCATGGGCTCCCTCAGGTTGGGCACA	508
QY	639	GCGTCTGGAGACAGCTGCGGGGTGCCCCCATATGCGTATCGATGCCAGAGSTATTAAAGGGG	698
Db	509	GCGTCTGGAGACCAAGCTGTGGGTCCCGCCACTACAGCTTCCCGGAGGTGATCCGGGGGG	568
QY	699	AAAAATATGATGGCCGCGCCGGGACAGATGTGGAGCTGTGGAGTCACTCTGTCGCCCTGC	758
Db	569	AGAAATATAGCAGCGCCGGAAGGCGAGAGTGTGGAGCTGCGCGTCACTCTGTCCGCTTGC	628
QY	759	TCGTGGGGGCTTGCCCTTWTGATGACGACAACTCCGCCAGCTGCTGGAGAGGTGAAC	818
QY	819	GGGGGCTCTCCACAGTGGCCCACTTATTCCTCCAAATTGGCCAGAGGCTCCTGAGGGGA	878
Db	689	GGGGGCTGTTCCACATGCGCCGCACTTATCCCGCCGCACTCCAGAGTCTGTGTACGGGGCA	748
QY	879	TGATCGAAGTGAAGCCCGGAAAAAGGCTAGTCTGGAGCAAAATTCAGAAAAATCCTTGGT	938
Db	749	TGAGGAGAGTGAACGCGCGGACCGCGGCTCAGCGTATAGACACATTCAAAACACATATGGT	808
QY	939	ACCTAAGCGGGGAACACAGGCCAGACCCGTGCTGTGAAGCCAGCCCTTGCGCCGCGGGTAG	998
Db	809	ATATATAGGGGGCAAGATGAGCCCGAAACC-----AGAGCAGCCCATTCCTCGCAAGGTGC	862
QY	999	CCATCGGAGGAGCTTGCATCCAAAGGAGAGCTGGAAACCCCGGACGTCCTAGAGAGCATTGGAT	1058
Db	863	AGATCCGCTCGCTGCGCCAGGCTTGGAGAGACATGCAACCCCGACGTCCTGACAGCATGTGACT	922
QY	1059	CACCTGGGCTGCTTACAGGGAGCCGGAAGAGCTGCATGCGAGCTGCGGACAGTGAAGAGAGA	1118
Db	923	CACCTGGGCTGCTTCCGAGAGACCGGAACAAGTGTGTGACAGACCTGCTCTCCGAGGAGAGAGA	982
QY	1119	ACCAAGAAAAGATGATATATTATCTGCTTTTGGATCGGAAGAGCGGATPACCAGCTGTG	1178

Db	983	ACCAGAGAGAAATGATTTACTTCTCTCTCTCTGAGCCGGAAGAAAGATACCCGAGCG	1042
OY	1179	AGGACCAAGAACTGCTCTCCCGGAAATGATGTGAACCCCCCGGAAGCGTGTGATCTC	1238
Db	1043	AGGATGAGGACCTGCCCCCCCCGGAACGAGATPAACCTCTCCCGGAACGTTGAGCTCC	11022
OY	1239	CGATGCTGAGCCGTCAACGGGAACGGCGACACAGCGGAGGATTCATGGAATCTGAGCA	12989
Db	1103	CGATGCTGAACGGCGACAGGGGAACGGGCGGCAABAGCAATCATGAGAGTGTCAAGC	1162
OY	1299	TCACCGATGCGGGGGGTGTGCTCCCTCTGTACCCACCGACGGGCTTGGAGATGCGCC	1358
Db	1163	TCAC-----GGACGGCGGTCTCCCGGTGCTGCGCGGGCCATTGAGATGGCC	1213
OY	1359	AGCACAGCCAGAGATCCCGTAGGTGAGTGGAGCCCTCAAGGGGTCTCTCCACGCGCTC	1418
Db	1214	AGCACAGCCAGAGATTCGGTCCATAGCGGTGCTCTTAAGGCTTTTCCACAGGCCAC	1273
OY	1419	TAAAGACCCCAAGAGATCGGGTTCCTTTCCTTACCCGGAACGGGGGCTTGGAGATAGG	1478
Db	1274	TCAGCAGCCCCG-----	1286
OY	1479	CTCGAGGGCGGGGCTCCCCGACTTCCAAAACGACAGCGCTGCTTCTCGGGGCCCCAGG	1558
Db	1287	-----	1286
OY	1539	GTTGGGGGCCCCGGGAGACAGCCCGCCCGGATGCGCCGCTCCACACCGCTCGCGGCG	1598
Db	1287	-----GTTGACCCCTACCCCTTCACCAAGGGGAGTCCCTCC	1326
OY	1599	CCCCAGGCTCCCGCGCTCTCTGCGGGGACCCCTTTCGACTGCTTCGACACGCCCC	1658
Db	1327	-----ACCCCAAGGGGAGACACTGTCCACACGCGCAA	1357
OY	1659	GGGCGCTGCCACCGGAGCCCCGGGGACACACACACCCCGCGGGGGTGGCGTGG	1718
Db	1358	AGGAGAGCCCGGCTGGCAGCGCCCAACCCACGCGCCCTCGAGCCCC-----AGCTCG	1411
OY	1719	GGGGAGCGCGTGGAGAGTGTCTCAACTCATTCGCAACAGGCTCTGGGCTCCCTC	1778
Db	1412	GAGGGGTGCCCTGGAGGGCGGCGCTCAACTCATTCGCAAGAGAGCTTCTGGGCTCACCC	1471
OY	1779	GCTTTCACCGGCGCAAGTGCAGAGTCCCTACCCCTGAGGAGATGTCCAGCTTGAAGCCAG	1838
Db	1472	GCTTCACCGCGGAACTGCAAGTTCCAGCGCGGGGAGATGTCCAACTGACACCG	1532
OY	1839	AGTCTCCCCCGAGCTGGCAAAACGCTTCTGGTGGGAATCTCATCTCTTGGACAAG	1898
Db	1532	AGTCTGCCCAAGAGCTGGCGCAAAAGTCTGTTGGGAATCTCATAGCCTGGAGAAAG	1592
OY	1899	AAGAACAAATATTCCTCTGTGCTTAAAGACAACCTCTCAGACGATCAAGACAGACATCG	1958
Db	1592	AGGAGAGATCTTCTGTGTCATCAAAAGACAACCTCTGACCTCATCAAGGCTGACATCG	1651
OY	1959	TCCATGCTTCTGTGATGCCAGCTGTGATCAAGTGTGCTGTCAAGACGAGCTTCA	2018
Db	1652	TGCAAGCCTTCTGTGATTTCCAGTCTCAAGACAGCGTATCTCCAAAGAGCTTCC	1711
OY	2019	GGGCGAGTACAAGGCGCAGTGGCGGCCCTCTCGTCTTCAAAAAGCCCGCTCGCTTCAGG	2078
Db	1712	GGGCGCAGTACAAGGCGCAGGCGGGGGCGACGCGTGTCCAAAGCCGGTCAAGTTCCAGG	1771
OY	2079	TGGACATCAGCTCTCTGTGAGGGTTCAGAGGCCCTCCCGGACGGGAGCGGACGGGAGGTG	2138
Db	1772	TGTGATTAACCTTACAGGAGGGT-----GGGAGGCGCGAGAAGAGA	1813
OY	2139	GTTGCACTACTCTGTCACCTTCACTCATCTCGGGTTCACGCGCTCGGTTCAAGGAG	2198
Db	1814	AGGCACTTACTCTGTCACCTTCAACCTGTCTCAAGGCCCAAGCGCTCGGTTCAAGAGG	1873
OY	2199	TGTGTGAGACATCCAGGCAAGCTCTCTGTGACATCAATGACAGCCCTCGGTGACGCCC	2258



Db 1524 GCTTCACCCGCGAAGCTGCAGGTCGACGCGGAGAGATGTCACACCTGACACAG 1583  
QY 1839 AGTCTCCCGGAGAGTGGCAAAAGCCTCTGGTTGGGAGCTTCATCTCTTGGACAAAG 1598  
Db 1584 AGTGTCTCCCGAGAGTGGCAAAAGCCTCTGGTTGGGAGCTTCATCTCTTGGACAAAG 1643  
QY 1899 AAGAACAAATATCTCTGGTGAAGGCAAAAGCCTCTGGAGCATCAAGAGAGATG 1558  
Db 1644 AGGAGCAGATCTTGTGTGTCATCAAGACAAAGCTGTGAGCTTCATCAAGAGCTGACATG 1703  
QY 1959 TCATGCTTCTGTGATCCCAAGCCTGTAGTCAAGTGTGTCATCAAGAGCTTCATCA 2018  
Db 1704 TGACGCTCTCTGTGATCCCAAGCCTGTAGTCAAGTGTGTCATCAAGAGCTTCATCA 1763  
QY 2019 GGGCGGAGTACAAAGGCGGAGTGGGCGGCTCTGCTTCCAAAAGCCGCTCCAGG 2078  
Db 1764 GGGCGGAGTACAAAGGCGGAGTGGGCGGCTCTGCTTCCAAAAGCCGCTCCAGG 1823  
QY 2079 TGACATGAGCTCTGTAGGAGTCCAGAGCCCTCCCGGAGGAGGAGGAGGAGG 2138  
Db 1824 TTGATATCTACCTACAGGAGGT-----GGGAGGCGGAGAGAGAG 1865  
QY 2139 GTGCACTCTACCTGCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2198  
Db 1866 ACAGCATCT 2195  
QY 2199 TGTGTGAGACATCCAGGACAGCTCTCTGAGCATCTATACAGCCCTCCGTCAGGCCC 2258  
Db 1926 TGTGTGAGACATCCAGGACAGCTCTCTGAGCATCTATACAGCCCTCCGTCAGGCCC 1985  
QY 2259 TGGCAGA 2265  
Db 1986 TGTGAGA 1992

RESULT 4  
US-10-054-579-3  
; Sequence 3, Application US/10054579  
; Patent No. US20020137913A1  
; GENERAL INFORMATION:  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Mathur, Brian  
; TITLE OF INVENTION: No. US20020137913A1 Human Kinases and Polynucleotides Encoding  
; FILE REFERENCE: LEX-0300-USA  
; CURRENT APPLICATION NUMBER: US/10/054, 579  
; PRIOR FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: US 60/263,378  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1827  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-10-054-579-3

Query Match 28.6%; Score 851.4; DB 12; Length 1827;  
Best Local Similarity 68.9%; Pred. No. 8.5e-190;  
Matches 1356; Conservative 0; Mismatches 416; Indels 195; Gaps 6;

QY 311 ATGAAGTGGAGCGGAGATGCGCATCTGGAAGCTCATGGAACACCAATGCTCCAG 370  
Db 1 ATGAAGTGGAGCGGAGATGCGCATCTGGAAGCTCATGGAACACCAATGCTCCAG 60  
QY 371 CTCACGAGCTCTAGAGAAACAAATATTTGTACTGTTCTGTGAGCAGCTCGGGG 430  
Db 61 CTCACGAGCTCTAGAGAAACAAATATTTGTACTGTTCTGTGAGCAGCTCGGGG 120  
QY 431 GGTGAGCTATTGACTTCTGTGTAAGAGGAGAGAGTACGCGCCAAAGAGCGCCGAAAG 490  
Db 121 GGTGAGCTATTGACTTCTGTGTAAGAGGAGAGAGTACGCGCCAAAGAGCGCGGAAAG 180  
QY 491 TTCTTCCCGCAGATGTGTGCTGCGCTGGACTTCTGCCACAGCTACTTCCTCCACAGA 550

Db 181 TTCTTCCCGCAGATGTGTGCTGCGCTGGACTTCTGCCACAGCTACTTCCTCCACAGG 240  
QY 551 GACCTAAAGCCCGAGAACCTGCTTTTGGATGAGAAAACATCCGATTCGAGATTC 610  
Db 241 GATCTGAAACCTGAAACCTCTCTGCTGGACGAGAAACAAACATCCGATTCGAGATTC 300  
QY 611 GGCATGAGCTCTCTGAGAGTGGGAGACAGCTCTCTGAGAGACAGTGGGAGTCCCTCAT 670  
Db 301 GGCATGAGCTCTCTGAGAGTGGGAGACAGCTCTCTGAGAGACAGTGGGAGTCCCTCAT 360  
QY 671 TATGCTGTCAGAGAGTATTAAGGGGAGAAATATGATGGCGCGGAGAGAGATGAG 730  
Db 361 TATGCTGTCAGAGAGTATTAAGGGGAGAAATATGATGGCGCGGAGAGAGATGAG 420  
QY 731 AGCTGTGAGTATCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 790  
Db 421 AGCTGTGAGTATCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
QY 791 CTCCGCGAGCTGTGAGAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 850  
Db 481 TTGCGACAGCTGTGAGAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
QY 851 CCAGATTCAGAGAGCTCTGAGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 910  
Db 541 CCCGATTCAGAGAGCTCTGAGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
QY 911 CTGAGCAAAATTCAGAAACATCTTGTGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 970  
Db 601 CTGAGCAAAATTCAGAAACATCTTGTGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 656  
QY 971 CTGAGCAAAATTCAGAAACATCTTGTGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1030  
Db 657 --AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 714  
QY 1031 GAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1090  
Db 715 GAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 774  
QY 1091 CATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1150  
Db 775 CTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 834  
QY 1151 GATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1210  
Db 835 GATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 894  
QY 1211 GAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1270  
Db 895 GAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 954  
QY 1271 GAGCGAAGTCCATGAGAGTCTCTGAGCATCAAGATGCGCGGAGGAGGAGGAGGAGGAG 1330  
Db 955 GAGCGAAGTCCATGAGAGTCTCTGAGCATCAAGATGCGCGGAGGAGGAGGAGGAGGAG 1005  
QY 1331 CCAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1390  
Db 1006 CCAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1065  
QY 1391 GCTTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1450  
Db 1066 GCTTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1107  
QY 1451 TCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1510  
Db 1108 ----- 1107  
QY 1511 CAGAGCTGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1570  
Db 1108 -----GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1127  
QY 1571 AGTGGCGGCTTCACAGCCCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1630

```

Db 1128 AAGGGCAGTCCCTCC-----ACC 1149
Qy 1631 CCCTTGACCTGCTCTGACACGCCCGGAGCAGTCCACCGGAGCCCGGGGACAA 1690
Db 1150 CCCAAGGGGACACCTGCTCCACAGCCAAAGAGAGCGGGGTGGACGCCCAACCCAG 1209
Qy 1691 CCACCCCCCAGCCCGGGGCGGTGGGAGCCCGCTGGAGAGTGTCTCAATCC 1750
Db 1210 CCCCCGTCAGCC-----AGCGTGGAGGGGTGCTGGAGGGCGCGGTCAATCC 1263
Qy 1751 ATCCGACAGCTTCTGGGGTCCCTGCTGATTCACCGGAGAGATGAGTCCCTACC 1810
Db 1264 ATCAAGACAGCTTCTGGGGTCCACCCGCTTCCACCGCGGAAATCAAGTCCAGAG 1323
Qy 1811 GCTGAGAGATGTCCAGCTTACAGCAGAGTCTCCCGGAGCTGGCAAAAGCTCTGG 1870
Db 1324 CCGAGAGAGATGTCCACAGTCCAGAGTGTCCCAAGAGCTGGCAAGAGTCTGG 1383
Qy 1871 TTGGGGAATCTTCTCTTGGAGCAAAAGAAACAATATCTCTGCTTAAGACAAA 1930
Db 1384 TTGGGGAATCTTCTTACAGCTTGGAGAGAGAGAGATCTTGTGTATCAAAAGAAA 1443
Qy 1931 CCTCTGACAGCATCAAGCAGAGCATGCTGCTTGTGTATCCGAGCTGAGT 1990
Db 1444 CCTCTGAGCTTCATCAAGGCTGAGCATGTGACGCTTCTGTGATTCAGTCTCAGC 1503
Qy 1991 CACAGTGTCTGTCTACAGCAGCTTCAAGGCGAGTACAAAGGCGAGCGCCCTCC 2050
Db 1504 CACAGGCTCATCTCCAAAGAGCTTCCGAGGCGAGTCAAGGCGAGCGGCGAGCG 1563
Qy 2051 GTCTTCCAAAAGCCCGCTTCCAGTGTGAGTCACTCTCTGAGGGTCCAGAGCC 2110
Db 1564 GTCTTCCAAAAGCCCGTCAAGTGTGAGTGTATCACTACAGGAGGT----- 1614
Qy 2111 TCCCGCGAGGAGCAGGAGGTGTGTATCTACTACTCCGTCACCTTCACTCATC 2170
Db 1615 -----GGGAGGGCGGAGAGAGAGAGAGAGCATCTACTCTTCACTCTGCTC 1665
Qy 2171 TCGGGTCCAGCCGCTGCTTCAAGCGAGTGTGAGAGCATTCAGGACCTCTCTGAGC 2230
Db 1666 TCGAGCCCGCCGCTCTTCAAGAGGAGTGTGAGAGCATTCAGGAGCCAGCTGTGAGC 1725
Qy 2231 ACATGAGACCGCTTCCGCTGAGCGCTTGGCAGACGAGAAAGCG 2277
Db 1726 ACACGAGACCGCTGCGCCGACGACTTGTGACACACCACTACTG 1772

RESULT 5
US-09-960-253-87
; Sequence 87, Application US/09960253
; Patent No. US20020123619A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Mohamath, Rodolph
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.556
; CURRENT APPLICATION NUMBER: US/09/960.253
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-960-253-87

Query Match 17.2%; Score 512; DB 10; Length 512;
Best Local Similarity 100.0%; Pred. No. 1.8e-110;
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 604 AGACTTCGGGATGGGCTGCGAGTGGGGGAGACAGCTCTCTGAGAGACACTGCGGGTC 663

```

```

Db 1 AGACTTCGGGATGGGCTGCGAGTGGGGGAGACAGCTCTCTGAGAGACACTGCGGGTC 60
Qy 664 CCCCCATTATGGCTGTCCAGAGGTATTAGGGGGAATAATATATATGCGCGCGGAGCA 723
Db 61 CCCCCATTATGGCTGTCCAGAGGTATTAGGGGGAATAATATATATGCGCGCGGAGCA 120
Qy 724 CATGTGAGCTGTGAGTCAATCTTCTGCGCTTGTCTGTGGGGGCTGTGCTTGTATGA 783
Db 121 CATGTGAGCTGTGAGTCAATCTTCTGCGCTTGTCTGTGGGGGCTGTGCTTGTATGA 180
Qy 784 CGAACACTTCCGACAGCTGTGAGAGTGAACGGGGGCTTTCACATGCCCCACTT 843
Db 181 CGAACACTTCCGACAGCTGTGAGAGTGAACGGGGGCTTTCACATGCCCCACTT 240
Qy 844 CATTCCTCCAGATTGCGAGAGCTCTCGAGGGGAATGATGAACTGGAGCCCGAAAAAG 903
Db 241 CATTCCTCCAGATTGCGAGAGCTCTCGAGGGGAATGATGAACTGGAGCCCGAAAAAG 300
Qy 904 GCTCAGTCTGGAGCAATTCAGAAACATCTTGTACTAGGCGGAGAAACAGAGCAGA 963
Db 301 GCTCAGTCTGGAGCAATTCAGAAACATCTTGTACTAGGCGGAGAAACAGAGCAGA 360
Qy 964 CCGCTGCTTGGAGCCAGCCCTGCGCGCGGTAGCCATGCGAGCTGCCATCAACGG 1023
Db 361 CCGCTGCTTGGAGCCAGCCCTGCGCGCGGTAGCCATGCGAGCTGCCATCAACGG 420
Qy 1024 AGAGTGGAGCCCGAGCTCTAGAGAGATGATCACTGGGCTGCTTCAAGGGAGCCGGA 1083
Db 421 AGAGTGGAGCCCGAGCTCTAGAGAGATGATCACTGGGCTGCTTCAAGGGAGCCGGA 480
Qy 1084 GAGGCTGATCGGAGCTGGCGAGTGAAGAG 1115
Db 481 GAGGCTGATCGGAGCTGGCGAGTGAAGAG 512

RESULT 6
US-10-102-558-1
; Sequence 1, Application US/10102558
; Publication No. US20030038706A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Ken-Shuo
; TITLE OF INVENTION: HUMAN PEN1B-RELATED GENE VARIANT ASSOCIATED WITH LUNG CANCERS
; FILE REFERENCE: 0 013922-3
; CURRENT APPLICATION NUMBER: US/10/102.558
; CURRENT FILING DATE: 2002-03-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2173
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(574)
; OTHER INFORMATION:
; US-10-102-558-1

Query Match 10.0%; Score 296.8; DB 9; Length 2173;
Best Local Similarity 73.4%; Pred. No. 9.1e-60;
Matches 401; Conservative 0; Mismatches 127; Indels 18; Gaps 1;
Qy 1695 CCCCCAGCCCGGCTGTGGGGGAGCCGCTGGAGAGTCTCTCAATCTCATCC 1754
Db 3 CCGCCCCGTCACAGCCCAAGGCTGAGAGGGGTGCTGAGAGGGCGGCTCACTCATCA 62
Qy 1755 GCAACAGCTTCTGGGCTCCCTGCTTTCACCGGCGCAAGATGACAGTCCCTACCGTC 1814
Db 63 AGAACAGCTTCTGGGCTCAACCCGCTTCCACCGCGGAAACTGCAAGTTCGAGCGCG 122
Qy 1815 AGAGATGTCCAGCTTCAAGCCAGAGTCTCCCGGAGCTGGGCAAAAGCTCTGGTTCG 1874
Db 123 AGAGATGTCCAGCTTCAAGCCAGAGTCTCCCGAGAGCTGGGCAAAAGTCTGGTTCG 182

```



QY	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300						
QY	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300						
Db	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110		
QY	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	
Db	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172
QY	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	
Db	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232
QY	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	
Db	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285							

RESULT 9  
 US-09-919-585-4  
 : Sequence 4, Application US/09919585  
 : Patent No. US20020115167A1  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Sun, Tian-Qiang  
 : APPLICANT: Feng, Jia-Jia  
 : APPLICANT: Reinhard, Christoph  
 : APPLICANT: Fanti, Wendy J.  
 : APPLICANT: Williams, Lewis T.  
 :  
 : TITLE OF INVENTION: ISOLATION OF DROSOPHILA AND HUMAN POLYNUCLEOTIDES ENCODING PAR-1  
 :  
 : TITLE OF INVENTION: POLYPEPTIDES ENCODED BY THE POLYNUCLEOTIDES AND METHODS UTILIZ  
 :  
 : FILE REFERENCE: PP-016093.002/200130.525  
 :  
 : CURRENT APPLICATION NUMBER: US/09/919,585  
 :  
 : NUMBER OF SEQ ID NOS: 22  
 :  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 :  
 : SEQ ID NO 4  
 :  
 : LENGTH: 2112  
 :  
 : TYPE: DNA  
 :  
 : ORGANISM: Homo sapiens  
 :  
 : US-09-919-585-4

OY	597	GCATTGCAGACTGTGGCATGGCGTCCTCCGACAGTGGAGGCGTCCCGACAGTGGAGGCGTCCCTCTGGAGACAGCATG	659
Db	488	AGATTCCAGACTTTTGGCTTCACGACATGATTCACCTTTTGGGAACAAGCTGGACACCTTCT	547
OY	657	GCGGGTCCCGCCATTATGCGTGTCCAGAGCTGATTAAGGGGGGAAAAATATGATGGCCGCC	718
Db	548	GTGGCAGTCCCGCTATGCTCGCCCOAGAACTTCCAGAGGGCAAAAATATGATGGACCG	607
OY	717	GGCGACATGTGGAGCTGTGGAGTCACTCTTCCGCCCTCTCTGTGGGGGCTCTGCGCT	778
Db	608	AGGTGGATGTGTGGAGCCCTAGGAGTTATCCCTATACACTGGTACGGCGATCCCTGCTT	667
OY	777	TTGATGACGACACACTCCGCCAGTGGCGTGGAGAGGTGAAGAAAGGGGGCGTCTCCACATGC	838
Db	668	TTGATGGACACAGACTCTAAGGAGCTGCGGGAGACGGGTATGAGGGGAAAAATACCGTATTC	727
OY	837	CCCACTTCATTCCTCCAGATTGCCAGAGCCCTCTGAGGGGAGATGATCAAGATGGAGCCCG	898
Db	728	CATTCATCATGTCCACGGAGCTGTGAAAACGTCTTAAGAAATTTCTCATTTCTTAATCCCA	787
OY	897	AAAAAAGCCTAGTCTGGAGCAAAATTCAGAAACATCTTGSTACTTAA-----GGCGGA	950
Db	788	GCAAGAGGGGCACTTTAAGCAAAATCAAGAAAGATCATGATGATGAATGTGGTCCACGAG	847
OY	951	AACAGAGCCAGACCCCGCTGGAGCGCAGCCGCTGGCGCGCGGGGTAGCCATGGGGA	1007
Db	848	ATGATGAACCTAAAGCTTTACCTGGAGGCCATCTCCCTGACTACAAAGACCCCGGGGGA	904

```

RESULT 10      US-09-919-585-7
; Sequence 7, Application US/09919585
; Patent No. US20020115167A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Tian-Qiang
; APPLICANT: Feng, Jia-Jia
; APPLICANT: Reinhard, Christoph
; APPLICANT: Fanti, Wendy J.
; APPLICANT: Williams, Lewis T.
; TITLE OF INVENTION: ISOLATION OF DROSOPHILA AND HUMAN POLYNUCLEOTIDES ENCODING PAR
; TITLE OF INVENTION: POLYPEPTIDES ENCODED BY THE POLYNUCLEOTIDES AND METHODS UTIL
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES
; FILE REFERENCE: PP-016093.002/200130.525
; CURRENT APPLICATION NUMBER: US/09/919,585
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2222
; TYPE: DNA
; ORGANISM: Homo sapiens
US-919-585-7

Query Match      8.1%; Score 240.6; DB 10; Length 2222;
Best Local Similarity 56.4%; Pred. No. 1,4e+46;
Matches 472; Conservative 0; Mismatches 319; Indels 6; Gaps 1;

```

QY 417 AGCAGCTCTGGGGGAGAGTCTATTCGACTACCTGTTAAAGAGGGGAGACTGACCCCA 476  
 Db 436 AGAAGCTAGTGGGAGAGGATTTGATTAACAGTGGCTCATGGAGAGTGAAGAA 495  
 QY 477 AGGAGGCCCCAAAGTCTTCCGCGAGATGTGTGCGCGCTGACTTGGCCAGACTACT 536  
 Db 496 AAGAGGCTCGAAGCAATTCGCGAGATAGTGTCTGCTGTGAGTACTGTCCAGAACT 555  
 QY 537 CCATTCGCCACAGACCTTAAGCCCGAGAACCTGCTTTGGATGAGAAAAACAACATCC 596  
 Db 556 TTAATGTCATAGAGCTTAAGAGCAGAAAACTGCTCTTGATGCTGATATGAAACATCA 615  
 QY 597 GCATGAGACTTGGGATGGCCCTCGACAGTGGGGGAGACGCTCCTGAGACCACT 656  
 Db 616 AGATTGAGAGCTTGGCTTACGCAATTAATTCACCTTGGGAGAACGCTGGACACCTTCT 675  
 QY 657 GCGGATCCCCCATTAATGCGTGCAGAGTGTATTAAGGGGAAAAATATGATGGCCGCG 716  
 Db 676 GTGGCAGTCCCTTATGCTGAGCCAGAACCTTCCAGGGGCAAAAATATGATGAGCCG 735  
 QY 717 GGGCAGACATGTGAGCTGTGAGTCACTCTTCGCTGCTGAGGCTGCTGCT 776  
 Db 736 AGGTGATGTGTGAGCTGAGAGTATACCTCTATACACTGTGTCAGCGGATCCGCTT 795  
 QY 777 TTGATGACGACAACTCCGCGACGCTGTGAGAGTGAAGGAGGCGCTCTTCCACATGC 836  
 Db 796 TTGATGACGACAACTCCGCGACGCTGTGAGAGTGAAGGAGGCGCTCTTCCACATGC 855  
 QY 837 CCCACTTATTCCTCCAGATTTGCCAGAGCCCTGAGGGGAGATGTCAGAGTGGAGCCG 896  
 Db 856 CATTTACATGTCACGAGACTGTAAACACCTGCTTAAGAAATTTCTTAAATCCCA 915  
 QY 897 AAAAAAGCTCAGTCTGAGCAAAATTCAGAAACCTCTTGTACCTAG-----GCGGA 950  
 Db 916 GCAAGAGAGGACCTTGTAGAGCAATTCATGAAGATGATGATGATGGGTCAAGAG 975  
 QY 951 AACACGAGCAGACCGCTGCTGAGAGCAGCCCTGCGCGCGGTAGCATGCGGA 1007  
 Db 976 ATGATGAACATAAGCCCTTACGTGTGAGGCCACCTGACTACAGAGACCCCGCGGA 1032  
  
 RESULT 11  
 US-09-938-842A-2657  
 ; Sequence 2657, Application US/09938842A  
 ; Patent No. US20020160378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Man9, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; FILE REFERENCE: SRIPI300-3  
 ; CURRENT APPLICATION NUMBER: US/09/938, 842A  
 ; PRIOR FILING DATE: 2001-08-24  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 3379  
 ; SEQ ID NO 2657  
 ; LENGTH: 1539  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-2657  
  
 Query Match 7.6%; Score 225.4; DB 9; Length 1539;  
 Best Local Similarity 56.9%; Pired. No. 4.3e-43;  
 Matches 434; Conservative 0; Mismatches 326; Indels 3; Gaps 1;

Db 55 TACAACTTGGGAGAACTCTGTATTGTTCTTCTTGTAGGGTGAAGATAGCTAGCAT 114  
 QY 245 TGCATCAGGGGTGAGAGGTCGCCATCAAGATCGTGAACCGGGAGAACCTGTGGA---G 301  
 Db 115 GCATTCAGAGGACCTAAGGTTGCTATCAAGATCCATCTCGCAAAATCAAGAACATG 174  
 QY 302 TCGGTGCTATGAAGTGTGAGCGGAGATGCCATCCCTGAAGCTCATCAACACCCACAT 361  
 Db 175 GAGATGAGAGGAAAGTGTGAGAGAGATCAAAATCTTGAACATATTATGATGCTTCAAC 234  
 QY 362 GTCTCAAGTCCACGAGCTGTACGAGAACAGAAATTTGTACCTGTCTGTGAGCAC 421  
 Db 235 ATCATCGTCTATGATGAGGTTATGAGACTCCACAGATATTATCTGTATGAGGAT 294  
 QY 422 GTCTCGGGGGTGTGACTATGCTGATGCTGTAAGAGGGGAGACCTACCCCAAGAG 481  
 Db 295 GTGAACTCTGTGAGCTATTTGATATATTTGTAAGAGGGTATGATGAGAGATGAG 354  
 QY 482 GCGGAAAGTCTCTCCGACAGATTTGTCTGCGCTGAGCTTGTCCACAGCTACTCATC 541  
 Db 355 GCGAGGAACCTTTTTCAGCAGATTAATTAAGAGAGTGGAAATCTGCCATCGAAACATGGTG 414  
 QY 542 TGCCACAGACCTTAAGCCCGAGAACCTGCTTTTGGATGAGAAACACATCCGATT 601  
 Db 415 GTTCACAGAGACCTCAAGCCCTGAAACCTTGTGACTTAAATGCAATGTAAGATT 474  
 QY 602 GCAAGCTTGGCATGGGCTGCCCTGACAGTGGGGAGACCCCTCCAGACCAAGCTGCGG 661  
 Db 475 GCTGATTTTGGCTGACGACACATATGAGATGTCTATTTTGAACAAGATGTGTG 534  
 QY 662 TCCCCCATTAATGCTGTCAGAGAGTATTAAGGGGAAAAATATGATGAGCCGCGGCA 721  
 Db 535 AGTCAAAATTAATGCGGCTCCAGAGTAAATTCGGGCAAGTATATGCGCCCTGAAGTA 594  
 QY 722 GACATGTGAGCTGTGAGATCATCTTGTGCGCTGCTGTGAGGGCTGTGCGCTTTAT 781  
 Db 595 GATGTGAGAGCTGTGTGTATCTTACCTTCTTCTGTGAGACCTTCAATTTAT 654  
 QY 782 GAGCAGAACCTCGGCCGAGCTGCTGAGAGTGAAGGAGGCGCTTCCACATGCCAC 841  
 Db 655 GATGAAACATTTCCCAACCTTTTAAGAAATTAAGGAGGAGATATACATTTACTAGC 714  
 QY 842 TTCAATCTCCAGATTCGAGAGCTCTGAGGGGAATGATGAGTGAAGGAGGAGG 901  
 Db 715 CATTTATCTCTGCTGTGTAGATTTGATCCCGGAGATGCTGTGATGACCCCATGAA 774  
 QY 902 AGCTCAGTGTGAGCAAAATTCAGAAACATCTTGTGACTAG 944  
 Db 775 CGAGTAAACATCTCTGAGATCCGCAACACCTTGTGCTCAAG 817  
  
 RESULT 12  
 US-09-815-915-3  
 ; Sequence 3, Application US/09815915  
 ; Patent No. US2002025931A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Millennium Pharmaceuticals, Inc.  
 ; APPLICANT: Meyers, Rachel  
 ; TITLE OF INVENTION: 3714, 16742, 23546, AND 13887 NOVEL  
 ; FILE REFERENCE: 38155-20006.00  
 ; CURRENT APPLICATION NUMBER: US/09/815,915  
 ; CURRENT FILING DATE: 2001-03-23  
 ; PRIOR APPLICATION NUMBER: US 60/191,846  
 ; PRIOR FILING DATE: 2000-03-24  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 2352  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-815-915-3

Query Match 7.4%; Score 221.4; DB 10; Length 2352;  
 Best Local Similarity 54.2%; Pred. No. 4.4e-42;  
 Matches 450; Conservative 0; Mismatches 381; Indels 0; Gaps 0;

```

QY 136 CCCCCACCCCCACCCCCACCCCCACCCCAATATGTTGGGCCCCCTATCGGCTGA 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30 CCCCCGGGGCCAGGTCAGGCGCCAGCAGAACCCCTCCGGGGTTTTCAGACATCGA 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 196 GAAGACGCTGGGCAAGAGACAGACAGGCGTGTAACTCGGGGTCACCTCATCAGGG 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 90 GCGGACCTCGGCAAGAGCAACTTCGGGGTGTGAAGTGGCGGGCATTCAGTACCAA 149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 TCAGAAAGTCCCATCAAGATCGTGAACCGGAGAGCTGTCGAGTGGTGTGATGAA 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 AACGACGGTTCATATAAATAATGATMAAACAGATTACATTCACCAATTGGAGAA 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 GGTGAGCGGAGATGCGCATTCCTGAGTCAACAGCACCACATGTCACCACTCCA 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 AATCTATCGTAGAGTTCAGCTGATGAAGCTTCTGAACCATCCACATCATTAACCTT 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 376 CGAGCTCTACGAGAACAAATATTTGTACTGTTCTGAGACACCTCTCGGGGGTGA 435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 CCAGCTTATGGAACAAGAGACATGCTTACATGTCACGTAATTTGCTAAATGAGA 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 436 GCTATTGACTACCTGTTAAAGAGGAGACGCGCCCAAGAGCGCCGAAGTCTT 495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 AATGTTGATTAATTTGCTTCCAACGGGACCTGATGAGAACAGGCGGGAAGATTT 389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 CCGCCAGATTTGCTGCTGCGGCTGCTGACAGCTACCTGCTGCGACAGAGACT 555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 CTGGCAATCTCTGCGCGCTGAGTCTGTCACGACCATCATCTGTCACCGGAGCT 449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 556 AAGCCCGAGACCTGCTTTTGTATGAGAAAACACATCCGATTGCGAGATTGCGCAT 615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 CAAGACCGAGACCTCTGCTGCTGATGGAACATGAGACATCAAGCTGGCAATTTTGA 509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 616 GCGCTCCCTGAGTGGGGGAGACCTCTGAGACACAGCTGCGGGGTCCTCCCATTA 675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 TGGGAATTTCTACAAAGTCAAGAGAGCCTCTGTCACGCTGTGGGAGACCCCTATG 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 676 GTGTCCAGAGTGTATTAAGGGGGAATAATATGATGCGCGGGGAGCATGTGAGGCTG 735
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 CCGCCCGAAGCTCTTGAAGGGAAGAGATGAAGGCCCCCACTGGACATCTGAGACT 629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 736 TGGAGTATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 795
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 630 GGGCGTGTGCTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 689
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 796 CCAGCTGCTGGAAGGTGAACGGGGCTTTCACATGCCCCCATTCATTCCTCCAGA 855
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 690 GACGCTGAGACAGCGGGTGTGAGAGGCGCTGTCGATCCCTCTTCATGCTCAAGA 749
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 856 TTGCGCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 915
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 750 CTGTGAGAGCTGTATCCGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 916 GCAAAATTCAGAAATCTCTTGTACTAGGCGGGAACACAGACCGACCC 966
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 810 CCAGATCCGCGACAGCGGTGATGCGGGCTGAGCCCTGCTTGCAGGAGCC 860
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13  
 US-09-815-915-1  
 ; Sequence 1, Application US/09815915  
 ; Patent No. US2002025931A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Millennium Pharmaceuticals, Inc.  
 ; TITLE OF INVENTION: 3714, 16742, 23546, AND 13887 NOVEL  
 ; TITLE OF INVENTION: PROTEIN KINASE MOLECULES AND USES THEREFOR  
 ; FILE REFERENCE: 38155-20006.00  
 ; CURRENT APPLICATION NUMBER: US/09/815,915

CURRENT FILING DATE: 2001-03-23  
 PRIOR APPLICATION NUMBER: US 60/191,846  
 PRIOR FILING DATE: 2000-03-24  
 NUMBER OF SEQ ID NOS: 19  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 1  
 LENGTH: 2968  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (140)..(2491)  
 NAME/KEY: misc\_feature  
 LOCATION: (1)..(2968)  
 OTHER INFORMATION: n = A,T,C or G  
 US-09-815-915-1

Query Match 7.4%; Score 221.4; DB 10; Length 2968;  
 Best Local Similarity 54.2%; Pred. No. 4.9e-42;  
 Matches 450; Conservative 0; Mismatches 381; Indels 0; Gaps 0;

```

QY 136 CCCCCACCCCCACCCCCACCCCCACCCCAATATGTTGGGCCCCCTATCGGCTGA 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 CCCCCGGGGCCAGGTCAGGCGCCAGCAGAACCCCTCCGGGGTTTTCAGACATCGA 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 196 GAAGACGCTGGGCAAGAGACAGACAGGCGTGTAACTCGGGGTCACCTCATCAGGG 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 GCGGACCTCGGCAAGAGCAACTTCGGGGTGTGAAGTGGCGGGCATTCAGTACCAA 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 TCAGAAAGTCCCATCAAGATCGTGAACCGGAGAGCTGTCGAGTGGTGTGATGAA 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 289 AACGACGGTTCATATAAATAATGATMAAACAGATTACATTCACCAATTGGAGAA 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 GGTGAGCGGAGATGCGCATTCCTGAGTCAACAGCACCACATGCTCAAGTCCA 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 AATCTATCGTAGAGTTCAGCTGATGAAGCTTCTGAACCATCCACATCATTAACCTT 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 376 CGAGCTCTACGAGAACAAATATTTGTACTGTTCTGAGACACCTCTCGGGGGTGA 435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 409 CCAGCTTATGGAACAAGAGACATGCTTTCATCTGCTGATTAATTTGTAAGAGA 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 436 GCTATTGACTACCTGTTAAAGAGGAGACAGCAGCCCAAGAGCGCCGAAGTCTT 495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 469 AATGTTGATTAATTTGCTTCCAACGGGACCTGATGAGAACAGCGCGGAGAAATT 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 CCGCCAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 529 CTGGCAATCTCTGCGCGCTGAGTCTGTCACGACCATCATCTGTCACCGGAGCT 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 556 AAGCCCGAAGACCTGCTTTTGTATGAGAAAACACATCCGATTGCGAGATTGCGCAT 615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 589 CAAGACCGAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 616 GCGCTCCCTGAGTGGGGGAGACGCTCTGAGAGACAGCTGCGGGGTCCTCCCATTTATG 675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 649 TGGGAATTTCTACAAATCTGAGAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 676 GTGTCCAGAGTGTATTAAGGGGGAATAATATGATGCGCGCGGAGACATGTGAGACTG 735
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 709 CCGCCCGAAGTCTTGAAGGGAAGAGATGAAGGCCCCCACTGAGCATCTGAGACT 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 736 TGGAGTATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 795
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 769 GGGCGTGTGCTGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 828
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 796 CCAGCTGCTGGAAGGTGAACGGGGCTTTCACATGCCCCCATTCATTCCTCCAGA 855
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 829 GACGCTGAGACAGCGGGTGTGAGAGGCGCTTCCTCCGATCCCTCTTCATGCTCAAGA 888
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 856 TTGCGCAGGCTCTGAGGGGATGATCGAAGTGAAGCCCGAAGAAAGCTCACTGGA 915
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 889 CTGTGAGAGCTGTATCCGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY 916 GCAATTCAGAACATCTGTGTACCTAGCGCGGAACACAGACGACCC 966  
DB 949 CCAGATCCGGAGACACCGGTGATCGGGCTGACCTCTTCCCGGAC 999

## RESULT 14

US-09-960-352-5086  
Sequence 5086, Application US/09960352  
Patent No. US20020137139A1  
GENERAL INFORMATION:  
APPLICANT: Warren, Wesley C.  
APPLICANT: Tao, Nenping C.  
APPLICANT: Byatt, John C.  
APPLICANT: Mathialagan, Nagappan  
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
FILE REFERENCE: 16511.006/37-21(10298)C  
CURRENT APPLICATION NUMBER: US/09/960,352  
NUMBER OF SEQ ID NOS: 15112  
SEQ ID NO 5086  
LENGTH: 275  
TYPE: DNA  
ORGANISM: Bos taurus  
OTHER INFORMATION: Clone ID: 22-LIB3058-032-Q1-K1-F5  
US-09-960-352-5086

Query Match 7.2%, Score 215; DB 10; Length 275;  
Best Local Similarity 90.2%, Pred. No. 5.6e-41;

Matches 230; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 941 CTAGCGGGAACACGACGACCCGCTGAGACCCGCGCGGGTAGCC 1000  
DB 1 CTGGGCGGGAAACGACGACCCGCTGAGACCCGCGCGGGTAGCC 60  
QY 1001 ATGGGAGCCCTGCATCCAGGAGAGGAGTGGAGCCGCTCTGAGAGCATG 1060  
DB 61 ATGGGAGCCCTGCATCCAGGAGAGGAGTGGAGCCGCTCTGAGAGCATG 120  
QY 1061 CTGGGCTCTTCAAGGAGCGGAGAGGCTGCATCCGAGCTCGAGTGGAGAGGAGAAC 1120  
DB 121 CTGGGCTCTTCAAGGAGCGGAGAGGCTGCATCCGAGCTCGAGTGGAGAGGAGAAC 180  
QY 1121 CAAAGAAAGATGATATATATCTGCTTTTGGATGAGAGAGGAGGATGCCAGTGTGAG 1180  
DB 181 CAAAGAAAGATGATATATATCTGCTTTTGGATGAGAGAGGAGGATGCCAGTGTGTG 240  
QY 1181 GACGAGACCTGCGCT 1195  
DB 241 GATCATGACTGTCT 255

## RESULT 15

US-09-799-875-4  
Sequence 4, Application US/09799875  
Patent No. US20020034780A1  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel  
APPLICANT: Kapeller-Libermann, Rosana  
TITLE OF INVENTION: No. US20020034780A1 Human Protein Kinases and Uses  
FILE REFERENCE: 35800/209996  
CURRENT APPLICATION NUMBER: US/09/799,875  
PRIOR FILING DATE: 2001-03-06  
PRIOR APPLICATION NUMBER: 60/182,059  
PRIOR FILING DATE: 2000-02-11  
PRIOR APPLICATION NUMBER: 09/659,287  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 5983

TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (115)...(3723)  
US-09-799-875-4

Query Match 6.9%, Score 207; DB 10; Length 5983;  
Best Local Similarity 53.6%, Pred. No. 1.6e-38;  
Matches 452; Conservative 0; Mismatches 390; Indels 1; Gaps 1;

QY 122 CCCGCTTACACCTTCCCTCCACCCACCCACCCACCCACCCACCCAAATGTGGCC 181  
DB 74 CCGGCTTACACCTTCCCTCCACCCACCCACCCACCCACCCACCCAAATGTGGCC 132  
QY 182 CCCATCGCTGGAGAAACGCTGGAGAGAGACAGACAGAGGCTGGTAACTCCGGGTC 241  
DB 133 TACTACGAGATCGACCGACCACTCGGCAAGGCACTTCGCGGTCAAGCGGCCACG 192  
QY 242 CACTGATCAGCGGGTCAGAGGTCCCATCAAGATCGTGAACCGGAGAGAGCTGTGAG 301  
DB 193 CACTGATCAGCGGGTCAGAGGTCCCATCAAGATCGTGAACCGGAGAGAGCTGTGAG 252  
QY 302 TCGGCTGATGAGAGGTGAGAGGAGATCCCATCTCTGAAGCTCATGCAACCCAT 361  
DB 253 GAAACTTGAAGAGATTTCCGGAAGTTCATTAATGATGATGCTTGCACCCCAT 312  
QY 362 GTCCCAAGCTCCACGAGCTCTAGAGAAACAAATATTTTACTCTGTGTGGAGC 421  
DB 313 ATCATCAGCTCTTACAGATTTATGAGACAGACGAGATTTATCTGTGACAGATAT 372  
QY 422 GTCTCGGGGGTGAAGCTATTCGACTTCTGTAAAGAGGAGAGAGTCCCAAGAG 481  
DB 373 GTCTCGGGGGTGAAGCTATTCGACTTCTGTAAAGAGGAGAGAGTCCCAAGAG 432  
QY 482 GCCGGAAGTCTTCCGCAAGATTTGTCTGCTGAGCTTGCACAGTACTTCATC 541  
DB 433 GCAGCTGGAGGATTCACACATCTCATAGCTGTCTTTTGTCTCTGTGGAACAT 492  
QY 542 TGCCAGAGAGCTTAAAGCCGAGAACCTGTTTGAATGAGAAACAAATCCCAT 601  
DB 493 GTTATCTGATTTAAAGCTGAAATTTACTTCTGAGTCCAAATCTGAATTCAAAT 552  
QY 602 GCAGACTTCGAGTGGCTTCCCTGAGGTGGGGGACAGCTCTCTGAGACAGCTGCGG 661  
DB 553 GCAATTTTGGTTTCAAGAACCTTCTTCACTCTGAGGAGAGCTCTGAAGACCTGGTGGC 612  
QY 662 TCCCTCCATTTGCTGTCAGAGGTGATTAAGGGGAAATATATGAGCGCGGGA 721  
DB 613 AGCCCTCCCTATGCTGACACTGAACTTTTGAAGGAAATATATGAGCGCGGA 672  
QY 722 GACATGTGAGCTGAGAGTATCTTCCCTGAGGTGGGGGAGTGGGCTGCTTGTAT 781  
DB 673 GACATGTGAGCTGAGAGTATCTTCCCTGAGGTGGGGGAGTGGGCTGCTTGTAT 732  
QY 782 GAGCACAACCTCCGACACTCTGAGAGAGTGAAGAGGAGGCTTCCATGAGCCAC 841  
DB 733 GAGCACAACCTCCGACACTCTGAGAGAGTGAAGAGGAGGCTTCCATGAGCCAC 792  
QY 842 TTCAATCTCCAGATTTGCAAGAGCTCTGAGAGGAGTGAAGTGAAGAGCCGGA 901  
DB 793 TTCAATCTCCAGATTTGCAAGAGCTCTGAGAGGAGTGAAGTGAAGAGCCGGA 852  
QY 902 AGGCTAGTGTGAGCAATTCAGAAATCTTGTGATGAGGAGGAGGAGGAGGAGG 961  
DB 853 AGGCTAGTGTGAGCAATTCAGAAATCTTGTGATGAGGAGGAGGAGGAGGAGG 912  
QY 962 GAG 964  
DB 913 AAC 915

Search completed: April 22, 2003, 03:03:27

Tue Apr 22 09:15:28 2003

us-10-003-690-1.rmpb

Page 12

Job time : 205.254 secs

---

1	966.6	32.4	2908	4	US-09-930-181-1	Sequence 1, Appl1
2	951	31.9	3364	3	US-09-930-181-3	Sequence 3, Appl1
3	206.6	6.9	1647	4	US-09-101-146-44	Sequence 44, Appl
4	200.2	6.7	1929	4	US-09-359-161-4	Sequence 4, Appl1
5	198.6	6.7	1742	4	US-08-557-006C-38	Sequence 38, Appl
6	198.6	6.7	2652	4	US-08-557-006C-39	Sequence 39, Appl
7	198.6	6.7	2761	4	US-08-557-006C-24	Sequence 24, Appl
8	182.6	6.1	1736	4	US-08-557-006C-37	Sequence 37, Appl
9	182.6	6.1	1783	4	US-08-557-006C-36	Sequence 36, Appl
10	177.8	6.0	2698	2	US-08-677-298-1	Sequence 1, Appl1
11	175.2	5.9	1747	4	US-08-557-006C-44	Sequence 44, Appl
12	145	4.9	213	4	US-09-930-181-5	Sequence 5, Appl1
13	135.6	4.6	2132	2	US-09-159-385-3	Sequence 3, Appl1
14	135.6	4.6	2132	4	US-09-186-277-3	Sequence 3, Appl1
15	130.6	4.4	1429	2	US-09-189-385-4	Sequence 4, Appl1
16	130.6	4.4	1429	4	US-09-186-277-4	Sequence 4, Appl1
17	116.6	3.9	1507	4	US-09-523-849-4	Sequence 4, Appl1
18	116.6	3.9	1544	4	US-09-734-673-1	Sequence 1, Appl1
19	116.6	3.9	2060	4	US-09-523-849-1	Sequence 1, Appl1
20	107.2	3.6	1599	3	US-09-256-665-1	Sequence 1, Appl1
21	107.2	3.6	1599	4	US-09-167-322-3	Sequence 1, Appl1
22	107	3.6	2610	2	US-09-212-771-1	Sequence 1, Appl1
23	107	3.6	2610	3	US-09-091-058-1	Sequence 1, Appl1
24	103	3.5	2549	4	US-09-467-082-3	Sequence 3, Appl1
25	102.4	3.4	1302	1	US-08-913-508-2	Sequence 2, Appl1
26	98.2	3.3	1244	4	US-08-755-758-1	Sequence 1, Appl1
27	98.2	3.3	1244	2	US-08-974-655-1	Sequence 1, Appl1

28	98.2	3.3	1244	4	US-09-283-011-1	Sequence 1, Appl
29	98.2	3.3	4162	2	US-08-459-446A-26	Sequence 26, Appl
30	98.2	3.3	4162	3	US-08-459-595A-26	Sequence 26, Appl
31	98.2	3.3	4162	3	US-08-459-505B-26	Sequence 26, Appl
32	98.2	3.3	4162	3	US-08-459-444-26	Sequence 26, Appl
33	98.2	3.3	4162	4	US-09-457-422-26	Sequence 26, Appl
34	98.2	3.3	4165	1	US-07-951-715A-26	Sequence 26, Appl
35	97.2	3.3	2104	4	US-09-613-930-1	Sequence 1, Appl
36	97	3.3	2061	4	US-09-800-960-1	Sequence 1, Appl
37	96.8	3.2	4335	2	US-08-631-097-3	Sequence 9, Appl
38	96.8	3.2	5886	4	US-08-810-712-9	Sequence 9, Appl
39	96.2	3.2	1282	2	US-08-878-989-12	Sequence 12, Appl
40	96.2	3.2	1282	4	US-09-272-796-12	Sequence 12, Appl
41	95	3.2	3972	4	US-09-770-170-1	Sequence 1, Appl
42	94.2	3.2	1466	2	US-08-749-902-2	Sequence 2, Appl
43	93.8	3.1	5328	4	US-09-428-711A-15	Sequence 15, Appl
44	93	3.1	1962	4	US-08-924-183-4	Sequence 4, Appl
45	93	3.1	1862	4	US-09-488-364-4	Sequence 4, Appl

## ALIGNMENTS

```

RESULT 1
US-09-930-181-1
; Sequence 1, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 V1
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(2112)
US-09-930-181-1

```

Query Match	32.4%;	Score 966.6;	DB 4;	Length 2908;
Best Local Similarity	70.1%;	Pred. No. 4.3e-192;		
Matches 1485;	Conservative	0;	Mismatches 439;	Indels 195; Gaps 6

QY	159	CCACACAGCCCAATATATGTGGGCCCATCGCTCGGAGAGAGACGTCGGCAAGGACAGA	21.8
Db	134	CGCACACGCGCAGATATTTTGGGCCCTTACCGGGCTGGAGAACCGCTGGCCAAAGGGGCGA	19.3
QY	219	CAGGCTGGTTAAATCGGGGTCACACTGCATCACGGGTCAGAAAGTCGCCATCAAGATCG	27.8
Db	134	CAGGTCTGTGTAAGTGGGGGTTCACTGCCCTGCACCTGCAGAAAGTGGCCATCAAGATCG	25.5
QY	279	TGAACCGGAGAAAGCTGTGGAGTCGGGTGATGAAGTGGAGCGGGAGATGCCATCC	3.8
Db	254	TCAACCGTAGAAGCTCAGCAGTCGGGTGTGATGAAGGTGGACGGGAGATCGCGATCC	31.3
QY	339	TGAAGCTATGAAACACCCCATGTCTCTCAAGCTCCACGAGCTTACGAGAAACAAGAT	39.8
Db	314	TGAACCTATTGAGCACCCCGACGCTCTAAAGCTGCACGACGTTATGAAACCAAAAAT	37.3
QY	399	ATTGTACCTGGTTCGTGAGCACGCTCGGGGGGTGAGCTATTGCATACCTGGTAAGA	45.8
Db	374	ATTGTACCTGGTGTAGAAACAGCTGTAGGTGGTGAAGCTTTTGACTACCTGGTGAAGA	43.3
QY	459	AGGGGAGATACAGCCCAAGGAGGCGCCAAAGTCTTCGCGACAGATTGTGTCTGCGCTGG	51.8
Db	434	AGGGGAGGCTACGCGCTAAGGAGGCGTGGAAAGTCTTCGCGCAATCATCTCTGCGCTGG	49.9
QY	519	ACTTCTGCCACAGCTACTTCATCTTGCACAGAGACCTAAAGCCCGAGAACCTGCTTTTG	57.8

```

Db 494 AGCTGTCACAGCAGCTCCATATGACACAGGATGTAAACCTGAAACCTCCGCTGG 553
OY 579 ATGAGAAAACAAACATCCGATTCGACACTTCGCGATGGCGTCCCTGGAGTGGGGACA 638
Db 554 ACAGAGAAACAAACATCCGATTCGACACTTCGCGATGGCGTCCCTGGAGTGGGGACA 613
OY 639 GCGTCTCGAGACAGCTGCGGGTCCCGCAATATGCGTTCAGAGGTATTAAGGGG 698
Db 614 GCGTCTCGAGACAGCTGCGGGTCCCGCAATATGCGTTCAGAGGTATTAAGGGG 673
OY 699 AAAAAATGATGGCGCGGCGACAGATGTGAGCTGTGAGTATCTCTTCGCGCTGC 758
Db 674 AGAAGTATGACGCGCGGAGAGGAGCTGTGAGCTGTGAGTATCTCTTCGCGCTGC 733
OY 759 TCGTGGGGGCTTCGCGCTTCGATGACAGCAACCTCCGACCTGCTGGAGAGTGAAC 818
Db 734 TGGTGGGGGCTTCGCGCTTCGATGACAGCAACCTCCGACCTGCTGGAGAGTGAAC 793
OY 819 GGGGCGTCTTCACATGCTCCCATCTTCTCCAGATTCGACAGCTCCGAGGGAGA 878
Db 794 GGGGCGTCTTCACATGCTCCCATCTTCTCCAGATTCGACAGCTCCGAGGGAGA 853
OY 879 TGATCGAATGAGAGCCGGAAGAAAGCTCAGTGTGAGCAAAATTCAGAAACATCTTGT 938
Db 854 TGATCGAATGAGAGCCGGAAGAAAGCTCAGTGTGAGCAAAATTCAGAAACATCTTGT 913
OY 939 ACCTAGCGGGGAAACACAGAGCCAGACCTGCTGAGAGCCAGCCCTGGCGCGGGTAG 998
Db 914 ATATAGGGGGCAAGAAATGAGACCCGAAACC-----AGAGCAACCCATTCCTCGAAGGTGC 967
OY 999 CCATGGGAGCCCTGCATTCACAGAGAGAGTGTGAGCCCGAGAGTCTTAAAGAGATGGAT 1058
Db 968 AGATCCGCTGCTGCTGCGAGCTGTGAGAGACATGACCCCGAGCTGTGAGAGAGTGCAT 1027
OY 1059 CACTGGGCTCTTCAGAGGAGCCGAGAGAGTGTGATGCGAGCTGGAGAGAGAGAGA 1118
Db 1028 CACTGGGCTCTTCAGAGGAGCCGAGAGAGTGTGATGCGAGCTGGAGAGAGAGAGA 1087
OY 1119 ACCAAGAAAAGATATATATATCTCTTTGATGATGAGAGAGAGAGAGAGAGTGTG 1178
Db 1088 ACCAAGAAAAGATATATATATCTCTTTGATGATGAGAGAGAGAGAGAGAGTGTG 1147
OY 1179 AGGACAGAGAGCTGCTCCCGGAAATGATGTGAGCCCGGAGAGAGTGTGATGATG 1238
Db 1148 AGGACAGAGAGCTGCTCCCGGAAATGATGTGAGCCCGGAGAGAGTGTGATGATG 1207
OY 1239 CCATGCTGAGCCGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1298
Db 1208 CGATGCTGAAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1267
OY 1299 TCACCATGCGGGGGTGTGCTCCCTGTACCCAGCCGAGGGGCTTGGAGATGGCC 1358
Db 1268 TGAC-----GAGAGGGGGTGTGCTCCCTGTACCCAGCCGAGGGGCTTGGAGATGGCC 1318
OY 1359 AGACAGCCAGAGATCCGAGAGTGTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1418
Db 1319 AGACAGCCAGAGATCCGAGAGTGTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1378
OY 1419 TAAGAGCCCAAGAGATGCGGTCTTCTTTCACCGAGAGCGGGGGCTGGAATGAGG 1478
Db 1379 TCAGCAGGCCCG----- 1391
OY 1479 CTCGAGGCGGGGGCTCCCGACTTCCAAAAGCAGAGAGCTGCTTTCGCGGGCCCAAGG 1538
Db 1392 ----- 1391
OY 1539 GTGGGGGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1598
Db 1392 -----GTGAGCCCTTCACCTTCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1431
OY 1599 CCCAAGGCTCCCGGCTCTGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1658

```

```

Db 1432 -----ACCCCAAGGGGAGACCTGTCCACAGCCAA 1462
OY 1659 GGGCAGATCCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1718
Db 1463 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1516
OY 1719 GGGGAGCGCGCTGAGAGAGTGTCTCAACTCCATCCGCAACAGCTTCTGGGCTCCCTC 1778
Db 1517 GAGGGGAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1576
OY 1779 GCTTTCACCGCGCAAGATGATGAGTGTCTTACCGGTGAGAGAGATGTCCAGCTTGACCCAG 1838
Db 1577 GCTTTCACCGCGCGCAAGATGATGAGTGTCTTACCGGTGAGAGAGATGTCCAGCTTGACCCAG 1836
OY 1839 AGTCTCTCCCGGAGAGTGTGCAAAAGCTGCTGAGTGTGCGGAATTCATCTCTTGGAGCAAG 1898
Db 1637 AGTCTCTCCCGGAGAGTGTGCAAAAGCTGCTGAGTGTGCGGAATTCATCTCTTGGAGCAAG 1696
OY 1899 AAGACAAATATTTCTGCTGTCTTAAAGCAAAACCTCTCAGAGCATCAAAAGCAGACATCG 1958
Db 1697 AGGAGCAGATCTTGTGATGATCAAAAGCAAAACCTCTGAGCTCATCAAGAGCTGACATCG 1756
OY 1959 TCCATGCTTCTGTGATGATCCCAAGCTGATGATGATGATGATGATGATGATGATGATG 2018
Db 1757 TCCATGCTTCTGTGATGATCCCAAGCTGATGATGATGATGATGATGATGATGATGATG 1816
OY 2019 GGGCCGAGTACAAAGGCGAGTGGGGGCGCTCCGCTTCTTCAAAAGCCGCTCGCTTCCAGG 2078
Db 1817 GGGCCGAGTACAAAGGCGAGTGGGGGCGCTCCGCTTCTTCAAAAGCCGCTCGCTTCCAGG 1876
OY 2079 TGACATCTGCTCTGTGAGAGTGTGAGAGCTTCCCGGAGGGAGAGAGAGAGAGAGTGTG 2138
Db 1877 TGATATCTACATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1918
OY 2139 GTGGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2198
Db 1919 AGGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1978
OY 2199 TGGTGGAGACATTCAGAGCAGCTCTGTGAGCACTCATGACAGCCCTCGTGGAGAGCC 2258
Db 1979 TGGTGGAGACATTCAGAGCAGCTCTGTGAGCACTCATGACAGCCCTCGTGGAGAGCC 2038
OY 2259 TGGCAGACGAGAAAGAGCG 2277
Db 2039 TGTGAGACACACTACTG 2057

RESULT 2
US-09-930-181-3
; Sequence 3, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 VI
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (482)..(2239)
US-09-930-181-3

Query Match 31.9%; Score 951; DB 4; Length 3364;
Best Local Similarity 69.9%; Pred. No. 7.9e-189;
Matches 1484; Conservative 0; Mismatches 440; Indels 199; Gaps 7;
OY 159 CCCAGACGCCCAATATGTGGGCCCTATGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 218

```

Db 257 CAGAGCAGCGGCTATGTTGGGCTTACCGGCTGGAGAAACGCTGGGCAAGGGGCGA 316  
 QY 219 CAGGGCTGTTAAATCTCGGGGTCCATGTCATCAGGGGTGAGAGGTCCATCAAGATCG 278  
 Db 317 CAGGCTGGTGAAGCTGGGGTTCATGCGTCACCTGCCAGAGGTGGCCATCAAGATCG 376  
 QY 279 TGAACCGGGGAAGCTGTGAGTGGTGTGATGAAAGTGGAGCGGAGATCGCATCG 338  
 Db 377 TCAACCGTGAAGCTGAGCGGTGGTGTGATGAAAGTGGAGCGGAGATCGGATCG 436  
 QY 339 TGAAGCTCATGAGAACCCACATGTCTCTCAAGCTCAGAGCTGTACGAGAACGAAT 398  
 Db 437 TGAAGCTCATGAGAACCCACATGTCTCTCAAGCTCAGAGCTGTATATAAACAAAT 496  
 QY 399 ATTT-----GTACCTGTGTTGAGACAGCTCGGGGGTGAAGTATTCAGTACCTGTA 454  
 Db 497 ATTTGATGAGTACTGGTGTAGAACAGCTGTCAAGTGGTGAAGCTTGTGACTACTGG 556  
 QY 455 AGAAGGGGAGCTGACGCCAGAGAGCGCGAAAGTTCTCCGCGAGATGTGTCTCG 514  
 Db 557 AAGAAGGGGAGCTGACGCCAGAGAGCTGTGAAAGTCTTCGCGAGATCATCTCTGCG 616  
 QY 515 CTGAGCTTGTGACAGCTACTCATCTGTCACAGAGACTAAAGCCGAGAACCTGCT 574  
 Db 617 CTGAGCTTGTGACAGCTACTCATCTGTCACAGAGACTGTAAACCTGAAACCTCTG 676  
 QY 575 TTGATGAGAAAAACAATCCGATTCGACACTTGGCATGGCGTCCCTGAGGTGGG 634  
 Db 677 CTGAGAGAGAAACAATCCGATTCGACACTTGGCATGGCGTCCCTGAGGTGGG 736  
 QY 635 GACAGCTCTGAGAGACAGCTGCGGGTCCCGCATATGCGTGTGACAGAGTGTAAAG 694  
 Db 737 GACAGCTCTGAGAGACAGCTGCGGGTCCCGCATATGCGTGTGACAGAGTGTAAAG 796  
 QY 695 GGGGAAAAATATGATGAGCGCGCGGAGACATGTGAGCTGTGAGTCACTCTTCCG 754  
 Db 797 GGGGAAAAATATGATGAGCGCGCGGAGACATGTGAGCTGTGAGTCACTCTTCCG 856  
 QY 755 CTGCTGTGGGGGCTGCTGCTTTGATGAGACAACCTCGCCAGCTGTGAGAGATG 814  
 Db 857 TTCTGTGTGGGGGCTGCTGCTTTGATGAGACAACCTCGCCAGCTGTGAGAGATG 916  
 QY 815 AAGCGGGGCTCTTCCATGAGCGCCCATCTCATCTCCAGATGTGCAAGAGCTCTGAG 874  
 Db 917 AAGCGGGGCTCTTCCATGAGCGCCCATCTCATCTCCAGATGTGCAAGAGCTCTGAG 976  
 QY 875 GGAATGATCGAAGTGAAGCGCGGAAAAAGCTCAAGTGTGAGCAATTCAGAAACATCT 934  
 Db 977 GGCATGATCGAAGTGAAGCGCGCGGAAAAAGCTCAAGTGTGAGCAATTCAGAAACAT 1036  
 QY 935 TGTACCTGAGCGGGGAAAAAGCGAGACCGGCTGTGAGCGGCGCCGCGG 994  
 Db 1037 TGTATATATAGGGGCAAGATGAGCGCGGACCC-----AGAGACCGCATCTCTGCAAG 1090  
 QY 995 GTAGCCATGGGAGCGCTGCCATCCAGCGAGAGAGTGGAGCGCGCGAGCTCTAGAGCAT 1054  
 Db 1091 GTAGCCATGGGAGCGCTGCCATCCAGCGAGAGAGTGGAGCGCGCGAGCTCTAGAGCAT 1150  
 QY 1055 GCATCACTGGGCTCTTCAAGGAGCGGAGAGGTGATCGAGAGCTGTGAGAGAGAG 1114  
 Db 1151 CACTCACTGGGCTCTTCAAGGAGCGGAGAGGTGATCGAGAGCTGTGAGAGAGAG 1210  
 QY 1115 GAGAAGCAAGAAAGATATATATATATCTGCTTTGATCGAGAGAGAGCGATCCAGC 1174  
 Db 1211 GAGAAGCAAGAAAGATATATATATATCTGCTTTGATCGAGAGAGAGCGATCCAGC 1270  
 QY 1175 TGTAGAGCAAGAGAGCGCTGCTCCCGGAAATGATGTGAGCGCGCGGAGAGGTGTGAT 1234  
 Db 1271 CAGAGAGATAGAGAGCGCTGCGCGCGGAAACGAGATAGACCTCTCCGGAAGGTGTGAGC 1330  
 QY 1235 TCTCCCTGTGAGCGCTGACGAGAGCGGAGCGGAGCGGAGAGTCAATGAGAGTCTG 1294

Db 1331 TCCCGATGCTGAACCGGCAAGCGGCAAGCGGCGGCGGCAAGCAATTCATGAGTGTCTC 1390  
 QY 1295 AGCATACCGAGTACCGGGGGTGTGTGCTCCCTGTACCCACCGAGCGGCTGTGAGATG 1354  
 Db 1391 AGCGTAC-----GGAGCGGGGCTCCCGGTGTGCTGTGCGCGGCGCATTTAGATG 1441  
 QY 1355 GCCCAGCAGCAGCAGAGATATCCGAGGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1414  
 Db 1442 GCCCAGCAGCAGCAGAGATATCCGAGGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1501  
 QY 1415 CCTTAAGCAGCCCAAGAGATCCGAGTCTTCTTACCGAGCGGCGGCGTGTGAGAT 1474  
 Db 1502 CCACTACAGCAGC----- 1514  
 QY 1475 GAGGCTGAGCGGGGGGCTCCCGACTTCCAAAGCAGAGCGTCTTGTGGGGGCC 1534  
 Db 1515 ----- 1514  
 QY 1535 AGGGGTGGGGGCGCGGGGAGAGAGAGCGCGCGCGCGCGCGCTCCACACCGCTGCGC 1594  
 Db 1515 -----CCCGGTGACCGCTTACCGCTTACCGAGAGGGGCGAGTCCCTCCG----- 1558  
 QY 1595 GGGCGCCAGAGCTCCCGCGCTCCTGTGGCGGAGCGCGCTTGTGACTGCTGTGACAG 1654  
 Db 1559 -----ACCCCAAGGGGAGACCTGTCCACAGC 1585  
 QY 1655 CCGCGGCGAGTCCACCGGAGCGCGGAGCGCGGAGCAACACCGCGCGCGCGGTGCG 1714  
 Db 1586 CCAAGGAGAGCGCGGCTGTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1639  
 QY 1715 GTGGGGGAGCGCGCTGTGAGAGAGTGTGATCACTTCACTCCGCAACAGTGTCTGAG 1774  
 Db 1640 GTGCGAGGGGTGCGCTGTGAGAGAGCGCGGCTGTGAGAGAGAGAGAGAGTGTGAG 1699  
 QY 1775 CTTGCTTTCAGCGGCGCGAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1834  
 Db 1700 CCGCGCTTCCAGCGCGCGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1759  
 QY 1835 CCAAGTCTCCCGGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1894  
 Db 1760 CCAAGTCTCCCGGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1819  
 QY 1895 AAGAAGCAAAATATCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1954  
 Db 1820 AAGAAGCAAAATATCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1879  
 QY 1955 ATGCTCATGCTTCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2014  
 Db 1880 ATGCTCATGCTTCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1939  
 QY 2015 TTTAGGCGGAGTACAAAGGCGAGTGTGAGGCGCGCGCGCTTGTGAGAGAGCGGCTG 2074  
 Db 1940 TTTAGGCGGAGTACAAAGGCGAGTGTGAGGCGCGCGCGCTTGTGAGAGAGCGGCTG 1999  
 QY 2075 CAGGTGAGTACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2134  
 Db 2000 CAGGTGAGTACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2041  
 QY 2135 GGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2194  
 Db 2042 GAGAGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2101  
 QY 2195 CAGGTGTGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2254  
 Db 2102 AGGGTGTGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2161  
 QY 2255 GCGCTGTGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2277  
 Db 2162 CACTTGTGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2184

RESULT 3  
 US-09-101-146-44

Query Match	6.94;	Score 206.6;	DB 3;	Length 1647;
Best Local Similarity	55.34;	Pred. No. 4e-34;		
Matches 423;	Conservative	0;	Mismatches 333;	Indels 3;
			Gaps	1
QY 177	TGGGCCCCCTATCGGCTGTGAGAGAAGACGCTGGGGCCAAAGACAGACAGGCGCTGGTTAAACTCG	236		
Db 38	TCGGCCACTATCATCTCGGGGACACGCTGGGGCTGGGACCTTCGGGAAAGTGAAGGTGG	97		
QY 237	GGGTCCACTGCATCAGCGGGTCAGAAGGTCCGCAACGATCCTGMAACCGGGGAAGACCTGT	296		
Db 98	GCAAGCAGCAATTGACTGTGACATTAAGTTGCTGTAAAGTACTCAACCGGCAAGAAATTTC	157		
QY 297	CGGAGTGGCTGTAT--GAAGGTGAGACGGAGATGCGCATCTCTGAAGTCATCGAAGC	353		
Db 158	GAACCTGTGACGTGGTGGGAAATCCGACAGAGATGCAGAACCTGAAAGCTTTTCAGGC	217		
QY 354	ACCCACATGTCCTCAAGTCCACAGACGTCACAGAAACAAGAAATATTGTACTGTGTC	413		
Db 218	ACCTCATATATATAAATGTACAGGTCATCAGTACACCGTCATATATTTCATAGTGTA	277		
QY 414	TGGAGCAGCTTCGGGGGGGTGAGCATTCGCACTACTCTGTTAAAGAGGGAGACTGCACG	473		
Db 278	TGGAAATATGTCTCAGAGAGAGAGCATTTTGAATTATATCTGTAAATATGSAAGTTGGAGG	337		
QY 474	CCAAAGAGGCCGAAAGTTCTTCGCCAGATTGTGTGCGCTGACATTGCCCACAGCT	533		
Db 338	AAAAGGAGAGTCGACGCTGTTCTCCAGCAGATCCCTTCTGGTGTGACATATTGTACAGCG	397		
QY 534	ACTGCATCTGCCACAGAGACCTAAAAGCCCGGAACCTGTTTGGATGGAAAAACAAGA	593		
Db 398	ATATAGTGTGCTCCAGAGAAATTGAAACCTGAAAGAGCTCTGCTTATGTGACACATGAAGG	457		

```

RESULT 4
US-09-359-161-4
Sequence 4, Application US/09359161A
Patent No. 6342656
GENERAL INFORMATION:
APPLICANT: Bradford, Kent J.
APPLICANT: Dahal, Peetambar
APPLICANT: Yang, Hong
APPLICANT: Cooley, Michael
APPLICANT: Downie, Bruce
APPLICANT: Gee, Oliver
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses
TITLE OF INVENTION: to Stress Conditions in Plants
FILE REFERENCE: 023070-093900US
CURRENT APPLICATION NUMBER: US/09/359,161A
CURRENT FILING DATE: 1999-07-21
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 1929
TYPE: DNA
ORGANISM: Lycopersicon esculentum
FEATURE:
OTHER INFORMATION: Lycopersicon esculentum plant homolog of yeast
OTHER INFORMATION: SNF1 kinase subunit of protein kinase (Ltsnfl)
US-09-359-161-4

```

	Query Match	6.7%;	Score 200.2;	DB 4;	length 1929;
	Best Local Similarity	54.4%;	Pred. No. 9,1e-33;		
	Matches	425;	Conservative	0;	Mismatches 353; Indels 3; Gaps 1
QY	167	GCCCAATATGTGCGCCCTATCGCGTGGAGAAAGCGTGGGCAGAAAGCAGACAGAGGGCTG	226		
Db	69	GACATATTTTACGGACATATTAACCTCGGGAAACACTTGGCATGTGATCGTTGGCANA	128		
QY	227	GTTAAACTCGGGGTCCATCGCATACGGGTGCAAGAGTCCCATCAAGATCGGAACCGG	286		
Db	129	GTTAAATAATGCTGAACATATACGTTAAACAGGGCCACAAAGTGTCTGTCAGATTTCTTAATCTT	188		
QY	287	GAGAAAGCTGTGCGAGTCGG--TGCTGATGAAGGTGAGCGGGAGATCCGCATCTCTGAAG	343		
Db	189	CGAAAAATCAGAAATATGAGACATGGAGGAGAAAGTCCGTAGAGAAATCAAAATATTGACA	248		
QY	344	CTCATCGGAACCCCAATGTCTCTCAAGCTCCACGAGCTTACGAGACAAGAATATTTTG	403		
Db	249	TTGTTCATGCAATCCCAATATTAACGGCTTTTGTAGGTCATAGAGACACCATCAGATATA	308		



```

; LENGTH: 2652
; TYPE: DNA
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1747)
; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with
; OTHER INFORMATION: nucleotide 1765
US-08-557-006C-39

```

```

Query Match      6.7%; Score 198.6; DB 4; Length 2652;
Best Local Similarity 54.6%; Pred. No. 2.2e-32;
Matches 418; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

```

```

OY 177 TGGGCCCCCTTGGGCTGAGAGAGCGTGGGCAAGAGACAGAGGGGCTGTTAACTCG 236
DB 45 TCGGACACTACGTGCTGGGGACACCCCTGGGCTGGGCACTTGGCAAGAGGATTTG 104
OY 237 GGGTCCACTGCATCAGCGGTGAGAGGTGCGCATCAAGATCGTGAACCGGAGAGCTGT 296
DB 105 GAGAACATCAATTGACGGCATTAAGTGGCAAGTTAATGATTAATGACAGAGATTTC 164
OY 297 CGGAGT---CGTGTGATGAGGTGAGCGGAGATCGCATCTGAAGCTCATCGAAC 353
DB 165 GCAGTTAGATGTTGTTGAAAAATTAACGAGAAATTCAAATCTTAACTCTTTCGTC 224
OY 354 ACCACATGTCCTCAAGCTCCAGAGCTCTACGAGAACAAATATTTGTAACCTGGTTC 413
DB 225 ATCTCTATTAATTAACAACCTACCAAGATGATGACACTCCACAGACTTTTATGTTAA 284
OY 414 TGGAGCAGCTCTCGGGGGGTGAGCTATTGCTACTCTGTGTAAGAGGAGAGACTGAGCC 473
DB 285 TGGAAATATGTTGTGAGAGGATTTGTTGCTGACTCATCTGTAATACAGGAGGTTGAAG 344
OY 474 CCAAGAGGCCCCAAGATTCTTCCGCAAGATTGTTCTGCGCTGAGCTTTCGCCAGCT 533
DB 345 AGGTGGAAGTCTGGCGCTCTTCAGAGAGATTGTTGCGGTGAGCTACCTGTACAGGC 404
OY 534 ACTCATCTCTGACAGAGACTTAAGCCGAGAACCTCTTTGGATGAGAAAAACAACA 593
DB 405 ACATGTTGTTCACAGGAGACTGAAGCAGAGAACGTGTTCTGTGAGCCGCCAGATGAATG 464
OY 594 TCCGATTCGACAGATTCGCGATGCGCTCCCTGCAAGTGGGGGAGACCTCTGAGACCA 653
DB 465 CTAAAGATGCTGACTCGGACTCTCTAATATGATGTCAGATGTAATTTCTACCACTA 524
OY 654 GGTGGGGTCCCCCATTTATCCCTGTCAGAGTGAATTAAGGGGAAAAATATGATGCC 713
DB 525 GGTGTGATGCGCAAAATTAATGACAGCCGAGGATCTCAAGAAAGCTGTATGCGGGTC 584
OY 714 GCGGGGAGACATGTGAGGCTGAGAGTCACTCTTCGCGCTCTGAGGGGCTGTC 773
DB 585 CTGAGTGTGATGTGAGAGCTGTGTTATCTCTGATGCCCCCTTCTGTGACACCTTC 644
OY 774 CTTTGTATGACAGCAACCTCCGCACTGCTGAGAGAGGTAAGGAGGCGCTCTTCACA 833
DB 645 CGTGTGAGATGAGACAGTGTGCTACGCTCTTAAGAGATCCGAGGAGGTGTTCTACA 704
OY 834 TGGCCCATCTTATCTCCAGATTTGCCAGAGCTCTCTAGAGGAGATGATCGAAGTGAAC 893
DB 705 TCCCGGAGTATCTCAACCGTTCTATTTGCCACTGTGATGATCAACATGTTGCGAGTGAAC 764
OY 894 CCGAAAAAGAGCTCAGTCTGAGCAAAATTCAGAAAATCATCTTGT 938
DB 765 CTTTGAAGCAGACACTATCAAGACATACGAGAGCATGAATGTT 809

```

```

RESULT 7
US-08-557-006C-24
; Sequence 24, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:

```

```

; APPLICANT: Berl, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forster, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NCAP/PHM3588/UST
; CURRENT APPLICATION NUMBER: US/08/557,006C
; PRIOR FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 24
; LENGTH: 2761
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA of rat
; OTHER INFORMATION: Liver AMP protein kinase
US-08-557-006C-24

```

```

Query Match      6.7%; Score 198.6; DB 4; Length 2761;
Best Local Similarity 54.6%; Pred. No. 2.2e-32;
Matches 418; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

```

```

OY 177 TGGGCCCCCTTGGGCTGAGAGAGCGTGGGCAAGAGACAGAGGGGCTGTTAACTCG 236
DB 61 TCGGACACTACGTGCTGGGGACACCCCTGGGCTGGGCACTTGGCAAGAGGATTTG 120
OY 237 GGGTCCACTGCATCAGCGGTGAGAGGTGCGCATCAAGATCGTGAACCGGAGAGCTGT 296
DB 121 GAGAACATCAATTGACGGCATTAAGTGGCAAGTTAATGATTAATGACAGAGATTTC 180
OY 297 CGGAGT---CGTGTGATGAGGTGAGCGGAGATCGCATCTGAAGCTCATCGAAC 353
DB 181 GCAGTTAGATGTTGTTGAAAAATTAACGAGAAATTCAAATCTTAACTCTTTCGTC 240
OY 354 ACCACATGTCCTCAAGCTCCAGAGCTCTACGAGAACAAATATTTGTAACCTGGTTC 413
DB 241 ATCTCTATTAATTAACAACCTACCAAGTATGACACTCCACAGACTTTTATGTTAA 300
OY 414 TGGAGCAGCTCTCGGGGGGTGAGCTATTGCTACTCTGTGTAAGAGGAGAGACTGAGCC 473
DB 301 TGGAAATATGTTGTGAGAGGATTTGTTGCTGACTCATCTGTAATACAGGAGGTTGAAG 360
OY 474 CCAAGAGGCCCCAAGATTCTTCCGCAAGATTGTTCTGCGCTGAGCTTTCGCCAGCT 533
DB 361 AGGTGGAAGTCTGGCGCTCTTCAGAGAGATTGTTGCGCGTGTGACTGTACAGGC 420
OY 534 ACTCATCTCTGACAGAGACTTAAGCCGAGAACCTCTTTGGATGAGAAAAACAACA 593
DB 421 ACATGTTGTTCACAGGAGACTGAAGCAGAGAGCTGTGCTGAGCCGCCAGATGAATG 480
OY 594 TCCGATTCGACAGATTCGCGATGCGCTCCCTGCAAGTGGGGGAGACGCTCTGAGACCA 653
DB 481 CTAAAGATGCTGACTTGGACTCTCTAATATGATGATGATGATGATGATTTCTAGAACTA 540
OY 654 GGTGGGGTCCCCCATTTATGCTGTCCAGAGTGAATTAAGGGGAAAAATATATGATGCC 713
DB 541 GGTGTGATGCGCAAAATTAATGACAGCCGAGGATCATCTCAGAAAGCTGTATCGGGTC 600
OY 714 GCGGGCAGACATGTGAGAGTGTGAGATCACTCTTCGCGCTCTGTTGGGGCTCTGC 773
DB 601 CTGAGTGTGATCTGAGAGTGTGAGTGTATCTGTATGCTGTCTGTGCGACCTTCC 660
OY 774 CTTTGTATGACAGCAACCTCCGCACTGCTGAGAGGTAAGGAGGCGCTCTTCACA 833
DB 661 CGTGTGAGATGAGCAGCTGCTCAAGCTCTTAAAGAGATCCGAGGGGTGTGTTCTACA 720
OY 834 TGGCCCATCTTATCTCTCAAGATTTGCCAGAGCTCTGTAGGGGAGATGATCGAAGTGAAC 893

```

Db 721 TCCCGAGTATCTCAACCGTTCTATTGCGCACTGCTGATCACAATGCTGGAGTGAGCC 780  
 QY 894 CCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAATCTTGGT 938  
 Db 781 CCTTGAAGGAGCACTATCAAGACATACAGAGCATGATGT 825

RESULT 8  
 US-08-557-006C-37  
 ; Sequence 37, Application US/08557006C  
 ; Patent No. 6258547  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berl, Rajindar K.  
 ; APPLICANT: Carling, David  
 ; APPLICANT: Forster, Robert A.  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
 ; FILE REFERENCE: NGAP/PM37588/JST  
 ; CURRENT APPLICATION NUMBER: US/08/557,006C  
 ; CURRENT FILING DATE: 1996-03-06  
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093  
 ; PRIOR FILING DATE: 1994-05-20  
 ; PRIOR APPLICATION NUMBER: GB 9310489.1  
 ; PRIOR FILING DATE: 1993-05-21  
 ; PRIOR APPLICATION NUMBER: GB 9318010.7  
 ; PRIOR FILING DATE: 1993-08-31  
 ; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 37  
 ; LENGTH: 1736  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Rat Liver AMP  
 ; OTHER INFORMATION: protein kinase  
 US-08-557-006C-37

Query Match 6.1%; Score 182.6; DB 4; Length 1736;  
 Best Local Similarity 53.3%; Pred. No. 4e-29;  
 Matches 408; Conservative 0; Mismatches 354; Indels 3; Gaps 1;

QY 177 TGGGCCCTTCGCGTGGAGAGCGCTGGGCAAGACAGACAGGGCTGTAACTCG 236  
 Db 38 TCGGACACTGCTGCTGGGAGACGCTGGGCTGGGCACTTGGCAAGAGATTG 97  
 QY 237 GGGTCCACTGCATCAGCGGTGAGAGGTCCGATCAAGATGCTGAACCGGAGAGCTGT 296  
 Db 98 GAAACATCATTAATGAGCGCATTAAGTGGCAATTAATCTTAATAGACAGAAATTC 157  
 QY 297 CGAGT---CGGTCTGATGAGGTGAGGAGGAGATCGCATCTGTAAGCTCATCGAAC 353  
 Db 158 GCAGTTAGATGTTGTGGAATAAATAACGAGAAATCAAAATCTAAACCTTTGCTC 217  
 QY 354 ACCACATGCTCCAGCTCCAGCGCTCTACGAGAACAGAAATATTGTACCTGGTTTC 413  
 Db 218 ATCTCTATTATTAACAATATACAGGTGATCAGCATCCACAGATTTTATGTGTA 277  
 QY 414 TGGAGCAGCTCTCGGGGGGTGAGCTATTGCACTGCTGTAAGAGAGGAGAGCTGACGC 473  
 Db 278 TGAATATGTGTGGAGTGAATTAATTTGACTACATCTGAGATGAGAGCGGTGAAG 337  
 QY 474 CCAAGAGGCGCCGAAATTTCTCCGCAAGATTGTCTGCGCTGAGACTTTCGACAGCT 533  
 Db 338 AGATGGAAGCAGGCGCTCTTTCAGCAGATTCTGTGCTGTGATTAATCTCATAGGC 397  
 QY 534 ACCGATCTGCTCCAGAGAGCTTAAGCCGAGACCTCTTTGGATGAGAAACAAACA 593  
 Db 398 ATATGTTGTTCACTGAGACCTGAAACAGAGAAATGCTCTGTGATCACAATGATG 457  
 QY 594 TCGCATTTGAGAGCTTGGAGATGCGCTCCCTGCAAGTGGGGGAGACGCTCTGAGACCA 653  
 Db 458 CCAAGATAGCGCAATTTGGGATTAATCTAATAGATGTGAGATGGAATTTCTGAGAACGA 517

QY 654 GCTCGGGTCCCCCATTAATGCTGTCCAGAGTGAATTAAGGGGAAAAATATGAGCC 713  
 Db 518 GTTCGGATCTCCAAATTAATGAGCAGCAGCTGAAATCATCTCAGAGATTTATGAGGTC 577  
 QY 714 GCGGGAGACATGTGGAGGTGAGGCAATCCCTTCGCGCTGCTGCTGGGGGTGCG 773  
 Db 578 CTGAAGTTGATATCTGAGCTGAGTGTGTTATCTGATGCTCTTCTTTGTTGGACCTCC 637  
 QY 774 CTTTGAAGCAGCAACCTCCGAGCTGTGAGAGAGTGAAGAGGGGCTCTTCCACA 833  
 Db 638 CATTTGATGAGAGCATGCTACGTACGTATTATTAAGAGATCGAGGGGTGCTTTTATA 697  
 QY 834 TCGCCCACTTCAATCTCCGAGATTGCGAGACCTCTCTGAGGGGATGATCAAGTGGAGC 893  
 Db 698 TCCAGATATCTCAATCGTGTGCTCCCACTCTCTGATCATATGCTCGCAGTTGACC 757  
 QY 894 CCGAAAAAGGCTCAGCTGAGCAAAATTCAGAAATCTTGGT 938  
 Db 758 CACTGAAGCAGCACTATCAAGACATTAAGAGAGCATGATGT 802

RESULT 9  
 US-08-557-006C-36  
 ; Sequence 36, Application US/08557006C  
 ; Patent No. 6258547  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berl, Rajindar K.  
 ; APPLICANT: Carling, David  
 ; APPLICANT: Forster, Robert A.  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
 ; FILE REFERENCE: NGAP/PM37588/JST  
 ; CURRENT APPLICATION NUMBER: US/08/557,006C  
 ; CURRENT FILING DATE: 1996-03-06  
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093  
 ; PRIOR FILING DATE: 1994-05-20  
 ; PRIOR APPLICATION NUMBER: GB 9310489.1  
 ; PRIOR FILING DATE: 1993-05-21  
 ; PRIOR APPLICATION NUMBER: GB 9318010.7  
 ; PRIOR FILING DATE: 1993-08-31  
 ; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 36  
 ; LENGTH: 1783  
 ; TYPE: DNA  
 ; ORGANISM: Human AMP protein kinase  
 US-08-557-006C-36

Query Match 6.1%; Score 182.6; DB 4; Length 1783;  
 Best Local Similarity 53.3%; Pred. No. 4.1e-29;  
 Matches 408; Conservative 0; Mismatches 354; Indels 3; Gaps 1;

QY 177 TGGGCCCTTCGCGTGGAGAGCGCTGGGCAAGACAGACAGGGCTGTAACTCG 236  
 Db 39 TCGGACACTGCTGCTGGGAGACGCTGGGCTGGGCACTTGGCAAGAGATTG 98  
 QY 237 GGGTCCACTGCATCAGCGGTGAGAGGTCCGATCAAGATGCTGAACCGGAGAGAGCTGT 296  
 Db 99 GAAACATCATTAATGAGCGCATTAAGTGGCAATTAATCTTAATAGACAGAAATTC 158  
 QY 297 CGAGT---CGGTCTGATGAGGTGAGGAGGAGATCGCATCTGTAAGCTCATCGAAC 353  
 Db 159 GCAGTTAGATGTTGTGGAATAAATAACGAGAAATCAAAATCTAAACCTTTGCTC 218  
 QY 354 ACCACATGCTCCAGCTCCAGCGCTCTACGAGAACAGAAATATTGTACCTGGTTTC 413  
 Db 219 ATCTCTATTATTAACAATATACAGGTGATCAGCATCCACAGATTTTATGTGTA 278  
 QY 414 TGGAGCAGCTCTCGGGGGGTGAGCTATTGCACTGCTGTAAGAGAGGAGAGCTGACGC 473  
 Db 279 TGAATATGTGTGGAGTGAATTAATTTGACTACATCTGTAACATGAGAGGGTTGAAG 338  
 QY 474 CCAAGAGGCGCCGAAATTTCTCCGCAAGATTGTCTGCGCTGAGACTTTCGACAGCT 533

```

Db 339 AGATGAGACGAGCGGCTTTCAGAGATTCGTGCTGTGATTTACTGTATAGC 398
QY 534 ACTCCATCTGCCACAGACCTAAAGCCGAGAACCTGCTTTTGATGAGAAAAACA 593
Db 399 ATATGCTGTTCATGACCTGAAACAGAGATGTCCTGTGATGACACATGATG 458
QY 594 TCCGATTCGACACTTCGCGATGCGCTCCTGCAAGTGGGGAGAACCTCTCGAGAC 653
Db 459 CCAAGATACCCGATTCGCGATTCATATATGATGATGATGATGATGATGATGAT 518
QY 654 GCTGCGGCTCCCGCATTCATGCTGTCAGAGGATTAAGGGGAAAAATATGATG 713
Db 519 GTTGGCGATTCGAAATTAATGACAGCACTGAAGCATCTCAGGAGATGATGAG 578
QY 714 GCGGCGACAGATGAGAGCTGAGATCACTCTTCCCTGCTGCTGCTGCTGCTG 773
Db 579 CTGAATGATATCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 638
QY 774 CATTGATGACGACAACTCCGCACTGCTGAGAGAGGATGAGAGGAGGCTCTT 833
Db 639 CATTGATGATGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 698
QY 834 TCCGCGATTCATTCCTCCGATTCGAGAGGCTCCTGAGGAGATGATGAGAG 893
Db 699 TCCGAGATATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 758
QY 894 CCGAAAAAGGCTCAGCTGAGCAAAATTCAGAAACATCTCTGCT 938
Db 759 CACTGAAGCAGCACTATCAAGACATAGAGAGATGATGATGATGATGATGAT 803

```

## RESULT 10

```

US-08-677-298-1
; Sequence 1, Application US/08677298
; Patent No. 5863729

```

## GENERAL INFORMATION:

```

APPLICANT: Pivnicka-Worms, Helen
TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN TKAC-1

```

```

NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:

```

```

ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder

```

```

STATE: CO
COUNTRY: USA

```

```

ZIP: 80303

```

## COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/677, 298

```

```

FILING DATE: 09-JUL-1996
CLASSIFICATION: 536

```

## ATTORNEY/AGENT INFORMATION:

```

NAME: Caruthers, Jennie M.
REGISTRATION NUMBER: 34,464

```

```

REFERENCE/DOCKET NUMBER: 9-96
TELECOMMUNICATION INFORMATION:

```

```

TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089

```

```

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

```

```

LENGTH: 2698 base pairs
TYPE: nucleic acid

```

```

STRANDEDNESS: double
TOPOLOGY: not relevant

```

```

MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:

```

```

NAME/KEY: CDS
LOCATION: 376..2565
US-08-677-298-1

```

## Query Match

```

Best Local Similarity 51.7%; Score 177.8; DB 2; Length 2698;
Matches 404; Conservative 0; Mismatches 377; Indels 0; Gaps 0;

```

```

QY 177 TGGGCCCCCTATGCGTGGAGAAAGCCCTGGCAAGACAGACAGAGGCTGTAACTCG 236
Db 533 TCGGAACCTAGACAGACTGTTGAAAACATGCGCAAGGGGATTTTGGAAAATTTGG 592
QY 237 GGGTCACCTGATCAGCGGTGAGAGGTGCGCATCAAGATCTGAAACCGGAGAGCTGT 296
Db 593 CAAGCATATCTCTTACAGCGAGAGAGGTGCAATTAATAATTAATTAATTAATTA 652
QY 297 CGAGATGCGTGTGATGAGAGGTGAGAGGAGAGATGCGATCCGAACCTCATGAAAC 356
Db 653 ATCCAAAGCTCTACAAAGCTCTTCAAGAGATGAAATGAAATGAAATGAAATGAA 712
QY 357 CACATGCTCTCAAGCTCCACAGCTCTACAGAGAAAGAAATATTTGATCTGCTG 416
Db 713 CCAATATAGTGAAGTATTTGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 772
QY 417 AGCAGCTCTCGGGGGGTGAGCTATTGACTACCTGTTAAAGGGGAGACTGACGCCA 476
Db 773 AATATGCAAGTGGAGGTGAAGTATTTGACTATTTGTTGTCACATGCGAGATGAAG 832
QY 477 AGGAGCCCGGAAAGTCTTCCGCGAGATGCTGCGGCTGCGGAGCTCTGCGACAG 536
Db 833 AAGAAAGCAAGATCTAATTTAGACAGATGCTGCTGAGTCAATATGCTACAGAAC 892
QY 537 CCATGTCACAGAGAGACCTTAAGCCCGAGAACCTGCTTTTGATGAGAAACATATCC 596
Db 893 GGATGCTACATGAGAGACCTCAAGCGTGAAGAAATCTATTTAGATGCGGATATGAC 952
QY 597 GCATTGCAACATTCGCGATGAGGCTCCCTCGAGGTGGGAGACCTCTCGAGACAGCT 656
Db 953 AATATGCAAGTGGAGGTGAAGTATTTGACTATTTGCGGCTGAACCTGACACAGTT 1012
QY 657 GCGGGTCCCGCATTTATGCGTGCAGAGGTGATTAAGGGGGAATAATAGTGGCCGC 716
Db 1013 GTGGCAGTCTCTCATAGAGACCTGAGCTCTTCCAGGGGCAAAATATGAGGGGCG 1072
QY 717 GGGCAGACATGAGAGCTGAGTGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 776
Db 1073 AAGTGATGATGAGAGCTGAGGCTGCTATTTATACACATGACATGAGGCTACCT 1132
QY 777 TTGATGACAGCAACCTCCGCGAGCTCTGAGAGAGTGAAGCGGGGCTTCCACATGC 836
Db 1133 TTGATGAGGCGCAAACTTAAGAACTGAGAGAGATTAAGAGGGGAAATACAGATT 1192
QY 837 CCCACTTCATTCCTCCAGATTTGCGAGAGGCTCTGAGGGGGAATGAGAGTGGAG 896
Db 1193 CCTTCTACTGCTTACAGACTGTGAAACCTTCTCAAGCTTCTGCTGCTTAAATCAA 1252
QY 897 AAAAAAGGCTCAGTGTGAGCAAAATTCAGAAACATCTTGTGCTTACAGGCGGAAAC 956
Db 1253 TTAAGCGGCGCACTGAGAGCAAAATCATGAAGCAAGGTGATCATTGACGGGCAAT 1312
QY 957 A 957
Db 1313 A 1313

```

## RESULT 11

```

US-08-557-006C-44
; Sequence 44, Application US/08557006C
; Patent No. 6258547

```

## GENERAL INFORMATION:

```

APPLICANT: Berl, Rajindar K.
APPLICANT: Forder, Robert A.

```

TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
FILE REFERENCE: NSAP/PM37588/USP  
CURRENT APPLICATION NUMBER: US/08/557,006C  
CURRENT FILING DATE: 1996-03-06  
PRIOR APPLICATION NUMBER: PCT/GB94/01093  
PRIOR FILING DATE: 1994-05-20  
PRIOR APPLICATION NUMBER: GB 9310489.1  
PRIOR FILING DATE: 1993-05-21  
PRIOR APPLICATION NUMBER: GB 9318010.7  
PRIOR FILING DATE: 1993-08-31  
NUMBER OF SEQ ID NOS: 44  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 44  
LENGTH: 1747  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: gene  
LOCATION: (1)..(1747)  
OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -  
OTHER INFORMATION: fragment begins at nucleotide 24 and ends with  
US-08-557-006C-44

Query Match 5.9%; Score 175.2; DB 4; Length 1747;  
Best Local Similarity 54.4%; Pred. No. 1.4e-27;

Matches 419; Conservative 0; Mismatches 343; Indels 8; Gaps 3;

QY 177 TGGGCCCTATCGCTGGAGAGACGCTGGGCAAGACAGAGCGCTGGTAACTCG 236  
DB 38 TCGGACACTACGCTGGGGGACACCTGGGCGCTGGGCAAAATGAAATTTG 97  
QY 237 GGGTCCATGCTACAGCGGTGACAGAGTCCGCAATCAAGATGTTGACCGGAAAGTGT 296  
DB 98 GAAATATCAATGAGAGCGCTAAAGTGGCAGTTAAGATCTTAAATGAGAGAGATTC 157  
QY 297 CGGAGT---CGTGTGATGAGGTGAGGTGAGATCGGCATCTGTAAGTCAAGCAAC 353  
DB 158 GCAGTTAGATGTTGTTGGAAAAATTAACGAGAAATTCAAATCTTAACTTTCGTC 217  
QY 354 ACCCATGTCCTCAAGCTCCAGACGCTTACAGAAACAAAGAAATTTTACCTGTTTC 413  
DB 218 ATCCATATATTAACAACCTTACCAAGTATGATCAGCAGCTCCAAAGACTTTTATGTTAA 277  
QY 414 TGGAGACGCTCTGGGGGGTGGAGCTATTCGACTACCTGTTAAAGAGGGGAGACTGACGC 473  
DB 278 TGGAAATATGCTGTGAGGTAATGTTGCACTATCTGTAACACGCGGAGGTTGAAG 337  
QY 474 CCAAGAGGCGCCGAAGATTCTCCGCAAGATTGTCGCTGGCTGAGCTTCG--CCACAG 531  
DB 338 AGGTGGAAGCTCCCGGCTTCCAGCAGATTTCTGTCCGCTGAGCTACTGTCTCACAG 397  
QY 532 CTACTCATCTGCCACAGAGACCTAAAGCCGAGAACCTGCTTTTGTGATGAGAAAAACAA 591  
DB 398 GCACATGTTGTCCACAGGAGCTGAAGCCAGAGAGAGTGTGTGAGCGCCAGATGAA 457  
QY 592 CATCCGATTCGAGACTTCGGCATGGCTCCCTTCAGAGTGGGGGACACCTTCCTGAGAC 651  
DB 458 TCGTAAGATAGCTGACTTCGACTCTCTTAATATGATGTCAAGATGTTCAACGAGAC 517  
QY 652 CAGCTCGGGTCCCGCATATATGCGTGTCCAGAGGTATTAAGGGGGAAAAATATATGAG 711  
DB 518 TACCTGTGATCCCAATATATGACAGACCGGAGGTATCTCAGAGAGCGTGTATGCGGG 577  
QY 712 CCGCCGCGGACAGATGAGAGCTGTGAGTCAATCTTTCGCTGCTGCTGAGGGGCTCT 771  
DB 578 TCTGTAGGTGATATCTCGAGCTGTGTATCTGTATCTCTCTCTCTGTGGACCT 637  
QY 772 GCGCTTTGATGAC---GACAACCTCCGCGAGCTCTGTGAGAGAGTGAACGGGGGCTTT 828  
DB 638 CCGCTTCAG 697  
QY 829 CCAGATGCCCCACTTATCTCTCCAGATGTCAGAGAGCTCTGAGGGGAAATGATGAGAT 888

DB 698 CTACATCCCGGAGATATCTCAACCTTTCTATTCACCTCTGCTGTGACATGCTGACAGT 757  
QY 889 GGAGCCCGGAAAAAGGCTCAGCTGTGAGAGCAAAATTCAGAAACATCTTGTGT 938  
DB 758 GGAGCCCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGATGAT 807

RESULT 12

US-09-930-181-5  
Sequence 5, Application US/09930181  
Patent No. 6455292

GENERAL INFORMATION:

APPLICANT: Origene Technologies  
TITLE OF INVENTION: Full-Length Serine Protein Kinase In Brain and Pancreas

FILE REFERENCE: 16U 101 VI  
CURRENT APPLICATION NUMBER: US/09/930,181

CURRENT FILING DATE: 2001-08-16  
NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn version 3.0  
SEQ ID NO 5

LENGTH: 213  
TYPE: DNA

ORGANISM: Homo sapiens  
FEATURE:

NAME/KEY: CDS  
LOCATION: (1)..(213)

US-09-930-181-5

Query Match 4.9%; Score 145; DB 4; Length 213;  
Best Local Similarity 86.5%; Pred. No. 1.4e-21;  
Matches 160; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 159 CCCAGCAGCCCAATATGTTGGGCCCTATCGGCTGGAGAGACGCTGGCAAGAGACAGA 218  
DB 29 CGCAGACGCGCAGTATGTTGGGCCCTACCGGCTGGAGAGAGAGAGAGAGAGAGAGAG 88  
QY 219 CAGGCTGTGTTAACTCGGGGCTCCATGCTACGAGGCTCAGAGAGTGGCCATCAAGATCG 278  
DB 89 CAGGTGTGGGAGAGCTGGGGGCTCAGCTGCTACCTCCAGAGAGTGGCCATCAAGATCG 148  
QY 279 TGAACCGGGAGAGCTGTGAGGCTGGTGTGATGAAGGTGAGAGAGAGAGAGAGAGAG 338  
DB 149 TCAACCGTGAAGAGCTCAGCAGAGTGGTGTGATGAAGGTGAGAGAGAGAGAGAGAG 208  
QY 339 TGAAG 343  
DB 209 TGAAG 213

RESULT 13

US-09-159-385-3  
Sequence 3, Application US/09159385  
Patent No. 5958748

GENERAL INFORMATION:

APPLICANT: AKIRA, SHIZUO  
TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE

FILE REFERENCE: PR-569  
CURRENT APPLICATION NUMBER: US/09/159,385

CURRENT FILING DATE: 1998-09-23  
EARLIER APPLICATION NUMBER: JP97/261589

NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3  
LENGTH: 2132

TYPE: DNA  
ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: CDS

LOCATION: (94)..(1455)

US-09-159-385-3

Query Match 4.6%; Score 135.6; DB 2; Length 2132;  
Best Local Similarity 54.1%; Pred. No. 2.6e-19;  
Matches 360; Conservative 0; Mismatches 279; Indels 27; Gaps 3;

```
QY 163 GCACGCCCAATATGTGGGCCCCATATGCGCTGGAGACAGCTGGGCAAGACAGACAGG 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 108 GCAGGAGAGAGCTGGAGACCATATGATGGGAGAGAGTGGGCAAGTTCATCAA 167
QY 223 GCGGTTAACTGGGGGTCCATGTCATCAGGGGTGAGAGGTGCGCATCAAGATGTGAA 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 168 GATCTGCGGAAGTCCCGGAGAGAGGAGGAGAGTACGACCAAGTTCATCAA 227
QY 283 CCGGGAAGAGCTGTC-----GAGTCGGTGTGATGAGAGTGAAGCGGAGAT 330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 228 GAAGCGCCCGCTGTCAATCCAGCGCGGTGGGTGAGCGGAGAGATCGAGCGGAGGT 287
QY 331 CGGCATCTGAACTGATCGAACAACCCATGTCCTCAAGCTCCAGAGCTTCACGAGAA 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 288 GAACATCTGCGGGAGATCCGGCACCCCAATCATCACTCCGACAGCATCTTCGAGAA 347
QY 391 CAAGAATATTTGACTGTTCTGAGACAGTCTGGGGGGGTGAGTATTCGACTACT 450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 348 GAAGAGGAGAGCTGCTCATCTCGAGCTGTCTGCGGGGAGCTTTGACTTCT 407
QY 451 GGTAAAGAGGAGAGCTGACGCCCAAGAGAGGCCGAAAGTTCTTCGCGCAGATTGTCTC 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 408 GCGGAGAAAGAGTGTCTCATCTCGAGCTGTCTGCGGGGAGCTTTGACTTCT 467
QY 511 TCGCGTGGAGTTCGCCACAGTACTCATCTGCCACAGAGACCTTAAGCCGAGACT 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 468 CCGCGTTCACCTGCTGACTCTTAAGCCGATCGACACTTTGAGTGAAGCCGAGAAAT 527
QY 571 GCTTTGATGAGAAAAAC-----AACATCCGATTCGAGACTTCGGGATGCG 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 528 CATGCTGTGGAACAAGACGTGCCCAACCAAGATCAAGCTCATGCTTCGGCATGCG 587
QY 619 GTCCCTGACAGTGGGGGACAGCTCTGAGACAGCTGCGGGTCCGCCATTTATGCTG 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 588 GCACAAGATGAGAGCGGGGAGAGTCAAGAACATCTTGGGCAAGCCGAGTTGTGGC 647
QY 679 TCCAGAGGTATTAAGGGGAAAAATATGATGGCGCGCGGCAACATGTGAGCTGTGG 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 648 CCGAGAGATTTGAACATATGAGCGGCTG---GGCTGAGAGCGGAGATGTGAGCATCG 704
QY 739 AGTCATCTCTTGGCCCTGCTGCTGGGGGCTCTGCTTGTATGACGACCACTCGGCA 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 705 TGTCAATCACTATATCTCTCTGAGCGGTGCAATCCCGTTCTTGCGGAGACCAAGCAGA 764
QY 799 GCTGCT 804
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 765 GACGCT 770
```

RESULT 14  
US-09-186-277-3  
; Sequence 3, Application US/09186277  
; Patent No. 6171841  
; GENERAL INFORMATION:  
; APPLICANT: AKIRA, SHIZUO  
; APPLICANT: KAWAI, TARO  
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE  
; FILE REFERENCE: 081356/0128  
; CURRENT APPLICATION NUMBER: US/09/186,277  
; EARLIER FILING DATE: 1998-11-05  
; EARLIER APPLICATION NUMBER: JP97/261589  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2132  
; TYPE: DNA  
; ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: CDS  
LOCATION: (94)...(1455)  
US-09-186-277-3

Query Match 4.6%; Score 135.6; DB 4; Length 2132;  
Best Local Similarity 54.1%; Pred. No. 2.6e-19;  
Matches 360; Conservative 0; Mismatches 279; Indels 27; Gaps 3;

```
QY 163 GCACGCCCAATATGTGGGCCCCATATGCGCTGGAGAACGCTGGGCAAGACAGACAGG 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 108 GCAGGAGAGAGCTGGAGACCATATGATGGGAGAGAGTGGGCAAGTTCATCAA 167
QY 223 GCGGTTAACTGGGGGTCCATGTCATCAGGGGTGAGAGGTGCGCATCAAGATGTGAA 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 168 GATCTGCGGAAGTGGCGGAGAGAGGAGGAGAGTACGACCAAGTTCATCAA 227
QY 283 CCGGGAAGAGCTGTC-----GAGTCGGTGTGATGAGAGTGAAGCGGAGAT 330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 228 GAAGCGCCCGCTGTCAATCCAGCGCGGTGGGTGAGCGGAGAGATCGAGCGGAGGT 287
QY 331 CGGCATCTGAACTGATCGAACAACCCATGTCCTCAAGCTCCAGAGCTTCACGAGAA 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 288 GAACATCTGCGGGAGATCCGGCACCCCAATCATCACTCCGACAGCATCTTCGAGAA 347
QY 391 CAAGAATATTTGACTGTTCTGAGACAGTCTGGGGGGGTGAGTATTCGACTACT 450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 348 GAAGAGGAGAGCTGCTCATCTCGAGCTGTCTGCGGGGAGCTTTGACTTCT 407
QY 451 GGTAAAGAGGAGAGCTGACGCCCAAGAGAGGCCGAAAGTTCTTCGCGCAGATTGTCTC 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 408 GCGGAGAAAGAGTGTCTCATCTCGAGCTGTCTGCGGGGAGCTTTGACTTCT 467
QY 511 TCGCGTGGAGTTCGCCACAGTACTCATCTGCCACAGAGACCTTAAGCCGAGACT 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 468 CCGCGTTCACCTGCTGACTCTTAAGCCGATCGACACTTTGAGTGAAGCCGAGAAAT 527
QY 571 GCTTTGATGAGAAAAAC-----AACATCCGATTCGAGACTTCGGGATGCG 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 528 CATGCTGTGGAACAAGACGTGCCCAACCAAGATCAAGCTCATGCTTCGGCATGCG 587
QY 619 GTCCCTGACAGTGGGGGACAGCTCTGAGACAGCTGCGGGTCCGCCATTTATGCTG 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 588 GCACAAGATGAGAGCGGGGAGAGTCAAGAACATCTTGGGCAAGCCGAGTTGTGGC 647
QY 679 TCCAGAGGTATTAAGGGGAAAAATATGATGGCGCGCGGAGACATGTGAGACTGTGG 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 648 CCGAGAGATTTGAACATATGAGCGGCTG---GGCTGAGAGCGGAGATGTGAGCATCG 704
QY 739 AGTCATCTCTTGGCCCTGCTGCTGGGGGCTCTGCTTGTATGACGACCACTCGGCA 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 705 TGTCAATCACTATATCTCTCTGAGCGGTGCAATCCCGTTCTTGCGGAGACCAAGCAGA 764
QY 799 GCTGCT 804
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 765 GACGCT 770
```

RESULT 15  
US-09-159-385-4  
; Sequence 4, Application US/09159385  
; Patent No. 5958748  
; GENERAL INFORMATION:  
; APPLICANT: AKIRA, SHIZUO  
; APPLICANT: KAWAI, TARO  
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE  
; FILE REFERENCE: PH-569  
; CURRENT APPLICATION NUMBER: US/09/159,385  
; EARLIER FILING DATE: 1998-09-23  
; EARLIER APPLICATION NUMBER: JP97/261589  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patentin Ver. 2.0

SEO ID NO 4  
LENGTH: 1429  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (10)..(1353)  
US-09-159-385-4

Query Match 4.4%; Score 130.6; DB 2; Length 1429;  
Best Local Similarity 56.5%; Pred. No. 2.5e-18;  
Matches 291; Conservative 0; Mismatches 209; Indels 15; Gaps 2;

QY 308 CTGATGAAGGTGGAGCGGGAGATGCCATCCTGAAGCTCATGAGACACCCACATGTCTC 367  
DB 181 CGGAGAGAGATCGAACCGAGGTGAGCATCTGCGCGAGATCCGCCACCCCAACATCATA 240  
QY 368 AACCTCCACGACGCTCTACGAGAACAGAAATATTGTACTGTCTGAGACGCTTCG 427  
DB 241 ACACCTGATGACGCTTCGAGAACAGACAGATGTGTCTGATCTGAGCTGTGTCTC 300  
QY 428 GGGGGTGAAGCTATTCGACTACCTGTTAAAGAGGGGAGACTGACGCCCAAGAGGCCGA 487  
DB 301 GGTGGCGAGCTTTTCGACTTCCTGCGCAGAGAGTCACTGACGAGAGATGAGGCCACG 360  
QY 488 AAGTTCTCCGCGAGATGTGTCTGCGCTGAGACTTCTGCCACAGCTACTCCATCTGCCAC 547  
DB 361 CAGTTCTCTAAACAAATCCTAGACGGTGTCCACTACCTGCACTCCAGCGCATCGCACAC 420  
QY 548 AGAGACCTAAAGCCCGAGAACCTGCTTTTGGATGAGAAAACAACATCCGC----- 598  
DB 421 TTGACCTGAAGCCCGAGAACATCATGTGTCTGAGACAAAGCAGCCAGCCGCCCATTT 480  
QY 599 ---ATTGCAGACTTCGGCATGGCGTCCCTGAGGTGGGGGACAGCCCTCTGAGAGACAGC 655  
DB 481 AAGCTATCGACTTTGGCATGCGCACAGAGATCGAGGCTGGCAGGAGTTCAAGAACATC 540  
QY 656 TGGGGTCCCGCCCATTTGGGTGTCCAGAGGTGATTAAGGGGAAAAATATGATGGCCGC 715  
DB 541 TTTGGCACACCCGAGTTGTGCGCCCGAGATGTAACATATGAGCCA---CTTGGCTTG 597  
QY 716 CGGGCAGACATGTGAGCTGTGAGATCATCTTTCGCCCTGCTGCTGGGGCTCTGCC 775  
DB 598 GAGGCTGACATGTGAGCATTTGGCTCATCATCTTCTGAGCGGAGCGTCCCA 657  
QY 776 TTTGATGACGACAACCTCGCGCAAGCTCTGAGAA 810  
DB 658 TTCTGGGCGAGACCAAGCAGAGACGCTGACGAA 692

Search completed: April 22, 2003, 00:25:56  
JOB time : 83.9763 secs





```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_67"
/issue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 207 a 246 c 274 g 160 t 2 others
ORIGIN

```

```

Query Match 25.5%; Score 759.8; DB 14; Length 889;
Best Local Similarity 98.5%; Pred. No. 7e-95;
Matches 788; Conservative 0; Mismatches 8; Indels 4; Gaps 2;

```

```

QY 189 GGGTGAAGAGAGGCGGGAAGAGAGAGAGGCGGTAACTCGGGGCTCCACTCA 248
DB 9 GGGTGAAGAGAGGCGGGAAGAGAGAGAGGCGGTAACTCGGGGCTCCACTCA 68
QY 249 TCACGGGTCGAAGAGTCGCCATCAAGATCGTGAACCGGAGAGAGTCGAGAGTCGTGC 308
DB 69 TCACGGGTCGAAGAGTCGCCATCAAGATCGTGAACCGGAGAGAGTCGTGCAGTCGTGC 128
QY 309 TGAATGAAGAGTCGAGAGAGTCGCCATCGTGAACCGGAGAGAGTCGTGCAGTCGTGC 368
DB 129 TGAATGAAGAGTCGAGAGAGTCGCCATCGTGAACCGGAGAGAGTCGTGCAGTCGTGC 188
QY 369 AGCTCCACGAGTCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 428
DB 189 AGCTCCACGAGTCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 248
QY 429 GGGGTGAGTCATTCGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 488
DB 249 GGGGTGAGTCATTCGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 308
QY 489 AGTTCCTCCGCGCAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 548
DB 309 AGTTCCTCCGCGCAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 368
QY 549 GAGACCTAAAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 608
DB 369 GAGACCTAAAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 428
QY 609 TCGGCAATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 668
DB 429 TCGGCAATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 488
QY 669 ATTATGCGTCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 728
DB 489 ATTATGCGTCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 548
QY 729 GAGAGTCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 788
DB 549 GAGAGTCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 608
QY 789 ACCTCGCGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 848
DB 609 ACCTCGCGCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 668
QY 849 CTCACAGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 908
DB 669 CTCACAGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 728
QY 909 GTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 965
DB 729 GTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 788
QY 966 CGTG-CCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 984
DB 789 CGTGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 808

```

```

RESULT 2
BOJ17903 881 bp mRNA linear EST 16-Jul-2002
LOCUS BOJ17903
DEFINITION AGENCOURT 8241190 lupski.symphathetic_trunk Homo sapiens cDNA clone
IMAGE:6186946 5', mRNA sequence.
ACCESSION BOJ17903
VERSION BOJ17903.1 GI:21856800
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 881)
AUTHORS NIH-MGC http://mgi.cni.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM13580 row: c column: 11
High quality sequence stop: 620.
Location/Qualifiers
1.881
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_67"
/issue_type="retinoblastoma"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCGCAGCGCCG-3' and
5'-GACACTCTTCAGATCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

```

```

BASE COUNT 192 a 260 c 271 g 155 t 3 others
ORIGIN
Query Match 24.3%; Score 723.6; DB 14; Length 881;
Best Local Similarity 95.3%; Pred. No. 6.2e-90;
Matches 755; Conservative 0; Mismatches 36; Indels 1; Gaps 1;
QY 570 TGGTTTGGATGAGAAAACAAATCCGATTCGAGATTCGCGATGAGGCTCCGTGAGC 629
DB 1 TGGTTTGGATGAGAAAACAAATCCGATTCGAGATTCGCGATGAGGCTCCGTGAGC 60
QY 630 TGGGGAGAGCGCTCGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 689
DB 61 TGGGGAGAGCGCTCGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 690 TTAAGGGGAAAAATATGATGCGCGCGGAGAGAGATGAGAGTCGAGTCATCTCT 749
DB 121 TTAAGGGGAAAAATATGATGCGCGCGGAGAGAGATGAGAGTCGAGTCATCTCT 180
QY 750 TCGCCCTGCTCGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 809
DB 181 TCGCCCTGCTCGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 810 AGGTGAAGCGGCGCTTCACATGCGCCACTTCATCTCCAGATTCGAGAGAGCTCC 869

```

Db 241 AGGTGAACGGGGGCTTCCACATGCCCACTTCATCTCCAGATTGCCAGACCTCC 300  
QY 870 TGAGGGGAATGATCGAAGTGGAGCCGAAAAAGGCTAGTCTGGAGCAATTCAGAAAC 929  
Db 301 TGAGGGGAATGATCGAAGTGGAGCCGAAAAAGGCTAGTCTGGAGCAATTCAGAAAC 360  
QY 930 ATCTGTGATACCTAGGCGGGAACACAGACGACCCGCTGCTGGAGCCGCTGGCC 989  
Db 361 ATCTGTGATACCTAGGCGGGAACACAGACGACCCGCTGCTGGAGCCGCTGGCC 420  
QY 990 GCGGGGATGACGAGCGAGCGCTGCTGATCCAGAGAGAGTGGAGCCGCTGCTGGAG 1049  
Db 421 GCGGGGATGACGAGCGAGCGCTGCTGATCCAGAGAGAGTGGAGCCGCTGCTGGAG 480  
QY 1050 GCATGGCATCAGTGGGCTGCTGAGGAGCGAGAGAGTGGAGCCGCTGCTGGAG 1109  
Db 481 GCATGGCATCAGTGGGCTGCTGAGGAGCGAGAGAGTGGAGCCGCTGCTGGAG 540  
QY 1110 AGGAGGAAACCAAGAAAGATGATATATATATCTGCTTTGGATCGAAGAGCGGTATC 1169  
Db 541 AGGAGGAAACCAAGAAAGATGATATATATATCTGCTTTGGATCGAAGAGCGGTATC 600  
QY 1170 CCAAGTGTGAGGACCAAGACCTGCTCCCGGAAATGATGTGACCCCGGCGG 1228  
Db 601 CCAAGTGTGAGGACCAAGACCTGCTCCCGGAAATGATGTGACCCCGGCGG 660  
QY 1229 GTGATATCTCCATGCTGAGCGCTGAGGAGGAGCGAGAGAGTGCATGGA 1288  
Db 661 GTGATATCTCCATGCTGAGCGCTGAGGAGGAGCGAGAGAGTGCATGGA 720  
QY 1289 GTGCTGACATCACCGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1348  
Db 721 GTGCTGACATCACCGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
QY 1349 GAGATGCGCCAG 1360  
Db 781 GCCTTGGCAAG 792

RESULT 3  
LOCUS BF529743 831 bp mRNA linear EST 11-DEC-2000  
DEFINITION 602044002F1 NCL\_CGAP\_Brn67 Homo sapiens cDNA clone IMAGE:4181688  
ACCESSION BF529743  
VERSION BF529743.1 GI:11617106  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 831)  
NIH-MGC http://mgi.ncl.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaaps\_r@mail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILNMI at:  
http://image.llnl.gov  
Plate: LLAM9495 row: c column: 01  
High quality sequence stop: 736.  
Location/Qualifiers  
1. 831

FEATURES  
Source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4181688"  
/clone\_id="NCL\_CGAP\_Brn67"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q

loss"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: brain; Vector: pCMV-Sport6; Site: 1; Not:  
Site: 2; Sali; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.3 kb. Constructed by Life  
Technologies. Note: this is a NCL\_CGAP library."  
BASE COUNT 185 a 237 c 258 g 151 t  
ORIGIN  
Query Match 21.9%; Score 651.2; DB 12; Length 831;  
Best Local Similarity 96.3%; Pred. No. 5e-80;  
Matches 709; Conservative 0; Mismatches 23; Indels 4; Gaps 4;  
QY 403 GTACTGTTCTGTGAGCAGCTCTGCGGGGGGTGAGCTATTCAGTACCTGTAAAGAGG 462  
Db 44 GTACTGTTCTGTGAGCAGCTCTGCGGGGGGTGAGCTATTCAGTACCTGTAAAGAGG 103  
QY 463 GAGACTGAGCGCCCA -GGAGCGCGGAAAGTTCTCCGCAATTTGTGCTGCTGACT 521  
Db 104 GAGACTGAGCGCCCAAGGAGAGCGCGGAAAGTTCTCCGCAAG -TGTGCTGCTGCTGACT 162  
QY 522 TCTGCCACAGCTACTCATCTGCGACAGAGACTTAAAGCCGAGAACTGCTTTGGATG 581  
Db 163 TCTGCCACAGCTACTCATCTGCGACAGAGACTTAAAGCCGAGAACTGCTTTGGATG 222  
QY 582 AGAAAAACAATCCGATTCAGACTTCGAGCTGCGCTCCCTGCGAGGTGGGGGAGAGCC 641  
Db 223 AGAAAAACAATCCGATTCAGACTTCGAGCTGCGCTCCCTGCGAGGTGGGGGAGAGCC 282  
QY 642 TCTGAGAGACAGCTGCGGGTCCCGCCATTTGCTGCTGCTGCTGCTGCTGCTGCTG 701  
Db 283 TCTGAGAGACAGCTGCGGGTCCCGCCATTTGCTGCTGCTGCTGCTGCTGCTGCTG 342  
QY 702 AATATGAGCGCGCGGCGAGAGTGTGAGCTGTGAGTATCTTCTGCGCTGCTGCTG 761  
Db 343 AATATGAGCGCGCGGCGAGAGTGTGAGCTGTGAGTATCTTCTGCGCTGCTGCTG 402  
QY 762 TGGGGGCTGCGCTTTGATGAGACAGAACTCCGCGAGCTCTGAGAGAGTGAACGGG 821  
Db 403 TGGGGGCTGCGCTTTGATGAGACAGAACTCCGCGAGCTCTGAGAGAGTGAACGGG 462  
QY 822 GCGCTTCCACATGCGCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 881  
Db 463 GCGCTTCCACATGCGCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 522  
QY 882 TCGAAGTGAAGCGCGGAAAGGCTCAGTGTGAGCAATTCAGAAATCCTTGTGACC 941  
Db 523 TCGAAGTGAAGCGCGGAAAGGCTCAGTGTGAGCAATTCAGAAATCCTTGTGACC 582  
QY 942 TAGGCGGAAACACAGAGCCAGCCGCTGAGAGCCAGCCGCTGAGAGAGTGAACCA 1001  
Db 583 TAGGCGGAAACACAGAGCCAGCCGCTGAGAGCCAGCCGCTGAGAGAGTGAACCA 642  
QY 1002 TGGGG -AGCTGCGCTCAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1060  
Db 643 TGGGGAAGCTGCGCTCAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 701  
QY 1061 CTGGGCTGCTTCAAGGAGCGGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1120  
Db 702 CTGGGCTGCTTCAAGGAGCGGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 761  
QY 1121 CAAGAAAAGATGATAT 1136  
Db 762 ACAAAATGTTTATAT 777

RESULT 4  
LOCUS BM675364/c 678 bp mRNA linear EST 27-FEB-2002  
DEFINITION UI-E-E10-ahr-p-12-0-UI-s1 UI-E-E10 Homo sapiens cDNA clone  
ACCESSION BM675364  
VERSION BM675364.1 GI:18985262

EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 678)  
Bonaldi,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics ([www.resgen.com](http://www.resgen.com)).  
The following repetitive elements were found in this cDNA  
sequence: 201-243 >(TAA)n#Simple\_repeat  
Seq primer: M13 Forward  
POLY-A=yes.

**FEATURES**

**SOURCE**

```

BASE COUNT
ORIGIN
120 a      200 c      223 g      134 t      1 others
TAG_SEQ=CCGCG*
TAG_LIB=UI-E-EJ0
TAG_TISSUE=human retina
optic nerve, CCATTAAGTC; retina, CCGCG; Retina Foveal and
Macular, GTCC; RPE and Choroid, ACCCA. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI).
/tissue_type="fetal eyes, lens, eye anterior segment,
/optic_nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="PH10B (Life Technologies) (T1 phage resistant)"
/notes="Organ: eye; Vector: pP773-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-EJ0 is a subtracted cDNA library constructed
according to Bonaldi, Lennon and Soares/ Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pP773-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes, AGAATCAGCA
; lens, CGAATTAGCGA; eye anterior segment, AATCGCGAT;

```

Query Match	20.9%	Score 623.4	DB 14	Length 678
Best Local Similarity	99.28%	Pred. No. 3,6e-76		
Matches 657; Conservative	0	Mismatches 2	Indels 3	Gaps 3
OY 2191	CAAGGAGGTGGGGAGACCATCCAGGACACAGCTCCGAGCAGCATGACCAAGCCCTTCGT	2250		
Db 678	CAAGGAGCTGGTGGAGACCATCCAGGACAGCTCCTGAGACTCATATGACCAAGCCCTTCGT	619		
OY 2251	GCAGGCCCTGTGGAGACGAGAAAGCAAGCGGGGCCAGACCCGGCCGTGTCCTCCCAACCCCG	2310		
Db 618	GCAGGCCCTGTGCAGACGACAGACACGCGGGCCCAAGACCCGGCCGTGTCCTCCCAACCCCG	559		

OY	2311	AAGCTGCAAGCCCCACCCCGGGCCGCACAGACCACGAGCTGAGCAGTGTCTCCCAGCGAG	2370
Dd	558	AAGCTTGCAAGCCCCACCCCGGGCCGCACAGACCACGAGCTGAGCAGTGTCTCCCAGCGAG	499
OY	2371	CCCCCCCAAGAACAAGACTCTTGGCCAACAAAGGGAGCCCTCTGCCCCGACCCACAG	2430
Dd	498	CCCCCCCAAGAACAAGACTCTTGGCCAACAAAGGGAGCCCTCTGCCCCGACCCACAG	439
OY	2431	GGCCGGGGAGGAGAGGGGACCCCCCTTCACACCCCTTCCGTGCCCCCAACTGTGAATCTG	2490
Dd	438	GGCCGGGGAGGAGAGGGGACCCCCCTTCACACCCCTTCCGTGCCCCCAACTGTGAATCTG	379
OY	2491	TAAATTAAGGCCCAAGAACATGTGCGGAGGGGGGGGAGACACAANAACGGGCTTGCCCTG	2550
Dd	378	T-AATTAAGGCCCAAGAACATGTGCGGAGGGGGGGGAGACACAANAACGGGCTTGCCCTG	320
OY	2551	CAGGGATGGGGCTCCACAGGCCGTGCCAACACTGGGGGGTGTTCTAGGGGAACAGGGGGCG	2610
Dd	319	CAGGGATGGGGCTCCACAGGCCGTGCCAACACTGGGGGGTGTTCTAGGGGAACAGGGGGCG	260
OY	2611	GGGAGCTGTTCTATTATTTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTC	2670
Dd	259	GGGAGCTGTTCTATTATTTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTC	200
OY	2671	TCTGGGGGTGGGGGAGGGGGAGGGAGCGGAGGTGGTGGGGTGGGCTTGCACAAATCCGAGC	2730
Dd	199	TCTGGGGGTGGGGGAGGGGGAGGGAGCGGAGGTGGTGGGGTGGGCTTGCACAAATCCGAGC	140
OY	2731	AGGGCCCTCTGCTCCCTGTGTGTGTCGCCAACCCCTCTTTCGCGGGCCCTGCTCCCTGGT	2790
Dd	139	AGGGCCCTCTGCTCCCTGTGTGTGTCGCCAACCCCTCTTTCGCGGGCCCTGCTCCCTGGT	80
OY	2791	CCTCCCCCAC- GACCTTCTGTAGCGAATTTGCTTCC- GGAAGAAATCTGGTTTTGCGGT	2848
Dd	79	CCTCCCCCACAGGACCTTCTGTAGCGAATTTGCTTCCGGAAGAAATCTGGTTTTGCGGT	20
OY	2849	GA 2850	
Dd	19	GA 18	

RESULT	5
LOCUS	BC819795
DEFINITION	BC819795 658 bp mRNA linear EST 22-MAY-2001 60278183Ft1 NCI-CGAP-Brn67 Homo sapiens cDNA clone IMAGE:4932680 5' mRNA sequence.
ACCESSION	BC819795
VERSION	BC819795.1 GI:14167382
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 658) NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
AUTHORS	Contact: Robert Strausberg, Ph.D.
JOURNAL	
COMMENT	

```

FEATURES
  source
    Email: c9apds-remail.nih.gov
    Tissue Procurement: David N. Louis, M.D.
    cDNA Library Preparation: Life Technologies, Inc.
    CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LMNL at:
    http://image.llnl.gov
    plate: L1AM10858 row: n column: 09
    High quality sequence stop: 612.
    location/Qualifiers
      1..658
        /organism="Homo sapiens"
        /db_xref="taxon:9606"

```







cloned directionally into pTX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TCAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., Program coordinator.

BASE COUNT 153 a 185 c 195 g 125 t  
ORIGIN

Query Match 18.1%; Score 540.4; DB 14; Length 658;  
Best Local Similarity 89.8%; Pred. No. 8e-65;  
Matches 591; Conservative 0; Mismatches 66; Indels 1; Gaps 1;

QY 820 GGGGCTTCCACATGCCCTTCATTCCTCCAGATGCCAGACCTCTCCGAGGAAAT 879  
DB 2 GGGGCTTCCACATGCCCTTCATTCCTCCAGATGCCAGACCTCTCCGAGGAAAT 61  
QY 880 GATCGAAGTGGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 939  
DB 62 GATCGAAGTGGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120  
QY 940 CTTAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 999  
DB 121 TCTGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
QY 1000 CATTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1059  
DB 181 CATTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240  
QY 1060 ACTGGGCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1119  
DB 241 TCTGGGCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
QY 1120 CCAAGAAAGATGATATATATATATATATATATATATATATATATATATAT 1179  
DB 301 CCAAGAAAGATGATATATATATATATATATATATATATATATATATATAT 360  
QY 1180 GACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1239  
DB 361 AGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
QY 1240 CATTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1299  
DB 421 CATTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
QY 1300 CACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1359  
DB 481 CACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
QY 1360 GACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1419  
DB 541 GACACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
QY 1420 AAGAGGCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1477  
DB 601 GAGAGGCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 658

RESULT 10  
AL538014 836 bp mRNA linear EST 13-FEB-2001  
LOCUS AL538014 L1L1\_F1013\_FBrn1 Homo sapiens cDNA clone CS0DF029YC21 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL538014  
VERSION AL538014.1 GI:12801507  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 836)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polyes, D.

TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr.  
Location/Qualifiers  
1. 836  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DF029YC21"  
/clone\_1lb="L1L1\_F1013\_FBrn1"  
/dev\_stage="pooled tissue from post conception fetuses (20  
week, 24 week and 26 week)"  
/lab\_host="DH10B"  
/note="Organ: Retal brain; Vector: pCMVSPORT 6; 1st strand  
cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-stranded cDNA was digested with Not I  
and cloned into the Not I and Eco RV sites of the  
pCMVSPORT 6 vector. Library was constructed by Life  
Technologies. Contact : Feng Liang Life Technologies, a  
division of Invitrogen 9800 Medical Center Drive Rockville  
Maryland 20850, USA Fax : (1) 301 610 8371 Email :  
filiang@life.com URL :  
http://fulllength.invitrogen.com"

FEATURES  
source

BASE COUNT 181 a 230 c 268 g 152 t 5 others  
ORIGIN

Query Match 17.7%; Score 527.6; DB 9; Length 836;  
Best Local Similarity 81.8%; Pred. No. 3.e-63;  
Matches 630; Conservative 4; Mismatches 138; Indels 2; Gaps 2;

QY 164 CACGCCCAATATGTGGGCCCCCTATCGGCTGAGAGAGAGCGTGGGCAAGAGACAGACAGG 223  
DB 65 CACGCCGAGATATGTGGGCCCCCTATCGGCTGAGAGAGAGCGTGGGCAAGAGAGT 124  
QY 224 CTGTTAACTCGGGGCTCACTGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 283  
DB 125 CTGTTAACTCGGGGCTCACTGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 184  
QY 284 CGGAGAGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 343  
DB 185 CGTGAAGAGGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 244  
QY 344 CTTATGAAACCCACATGCTCTCAAGCTCCACAGAGCTTCAAGAGAAAGAAATATTTG 403  
DB 245 CTTATGAAACCCACATGCTCTCAAGCTCCACAGAGCTTCAAGAGAAAGAAATATTTG 304  
QY 404 TACCTGTTTGAAGACAGCTCTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 463  
DB 305 TACCTGTTTGAAGACAGCTCTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 364  
QY 464 AGACTAGCGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 523  
DB 365 AGCTTAGCGCTTAAGAGGCTTCGAGAGTCTTCGAGAGATCATCTGCGCTGACTTC 424  
QY 524 TGGCAGAGTACTCATCTGACACAGAGACCTTAAGGCGGAGAGAGCTTTTGGATGAG 583  
DB 425 TGGCAGAGCAGCTCAATATGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 484  
QY 584 AAAAACAATCCGATTCGAGACTTGGCATGGCGCTCTGACAGTGGGGGAGACGCTC 643  
DB 485 AAGAAACAATCCGATTCGAGACTTGGCATGGCGCTCTGACAGTGGGGGAGACGCTC 544  
QY 644 CTGAGACAGAGCTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 703  
DB 545 TTGGAACACAGCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 604  
QY 704 TATGATGCGCGCGGAGACATGTGAGAGCTGTGAGATCACTCTTCCGCTGCTGCTG 763  
DB 605 TATGAGAGGCGG-GAASGAGAGCTGTGAGAGCTGTGAGATCACTCTTCCGCTGCTGCTG 663  
QY 764 GGGGCTTCCCTTTGATGAGCAACACTTCGCGAGCTGTGAGAGAGTGAAGAGGGGGC 823

D <sub>b</sub>	664	GGGGCTCTGCCCCCTTGGAGATGACAACTTGGCACA - CTGCTGGAAAGGTGCAAGCGGGGC	722
Q <sub>y</sub>	824	GTCCTCCACATGCGCCCATCTTATTCCTCCAGATTTGCCAGAGCGCTCCTTAGGGGAATGATC	883
D <sub>b</sub>	723	GTGTTCACATGCGCCCATCTTATTCGCCGCSACTGCCAGATCTCTACGGGGCATGATC	782
Q <sub>y</sub>	884	GAAGTGCAGCCCGAAAAAAGGCTCAGTCTGGAGCAAAATTCACAAACATTCCTGG	937
D <sub>b</sub>	783	GAGGTGAGCGCCGACGCGGCTCAGCGTAGAGCAATTCAGAAAAACATATATG	836

RESULT 11	BF529630	936 bp	mRNA	linear	EST 11-DEC-2000
LOCUS	BF529630				
DEFINITION	602048585F1 NCI CGAP_Brn67 Homo sapiens CDNA clone	5', mRNA sequence.			IMAGE:4181581
ACCESSION	BF529630				
VERSION	BF529630.1	GI:11616993			
KEYWORDS	EST.				
SOURCE	human.				

REFERENCE	1 (bases 1 to 936)
AUTHORS	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Tissue Procurement: David N. Louis, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINTL at:  
<http://image.llnl.gov>  
Plate: LLAM9494 row: n column: 14  
High quality sequence start: 8  
High quality sequence stop: 568.

FEATURES	Location/Qualifiers
source	1. .936

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4181581"
/clone_id="NCL_CGAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with ip/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCL_CGAP Library."

```

Query Match	17.5%;	Score 522.4;	DB 12;	Length 936;
Best Local Similarity	97.6%;	Pred. No. 1.8e-62;		
Matches	562;	Conservative	0;	Mismatches 11; Indels 3; Gaps
QY	986	GGCCCGCCGGTAGCCATCCGAGGCGCTGCCATCCACGAGAGAGCTGGACCCCGACGCTCTTA	1045	
Db	5	GGCCCGCCGGTAGCCATCCGAGGCGCTGCCATCCACGAGAGAGCTGGACCCCGACGCTCTTA	64	
QY	1046	GAGAGCATGCGATACACGTGGCGCTTCAAGGACCGGAGAGGCTGCATCGGAGACTCGGC	1105	
Db	65	GAGAGCATGCGATACACGTGGCGCTTCAAGGACCGGAGAGGCTGCATCGGAGACTCGGC	124	
QY	1106	AGTAGGAGAGGAACCAAGAAATGATATATATATCTGCTTTTGGATCGGAAGAGACGG	1165	
Db	125	AGTAGGAGAGGAACCAAGAAATGATATATATATCTGCTTTTGGATCGGAAGAGACGG	183	
QY	1166	TATCCCACTGTGTGAGAACCAAGACCTGCTCCCGGAATGATGTTCAGACCCGCCCGGAAG	1225	

Db 184 TATCCACGCTGTGAGGACCAAGACACTGCTCCCGGAATGATGTGTAACCCCCCGGAAG 243

QY 1226 CGTGTGGATTTCCTCCATGCTAGACCGCTACGGGAAGCGGCGACACAGAGCGGAATCCATG 1285

Db 244 CGTGTGGATTTCCTCCATGCTAGACCGCTACGGGAAGCGGCGA - CAGAGCGGAATCCATG 302

QY 1286 GAAGTCTCGACATCACCGAATGCCGGGGGTGTGGTCTCCCTGTATCCACCCGACGGGCC 1345

Db 303 GAAGTCTCGACATCACCGAATGCCGGGGGTGTGGTCTCCCTGTATCCACCCGACGGGCC 362

QY 1346 TTGGAGATGGGCCAGACAGACAGATCCGCTAGACGCTAGTGGAGCCCTCCACGGCTGTG 1405

Db 363 TTGGAGATGGGCCAGACAG - CAGAGATCCGCTAGACGCTAGTGGAGCCCTCCACGGCTGTG 421

QY 1406 TCTCTCAGCCCTCTAAGCAGGCCCAAGAGATCCGGTCTTTTCTTTTCCCGAGCCGGGG 1465

Db 422 TCTCTCAGCCCTCTAAGCAGGCCCAAGAGATCCGGTCTTTTCTTTTCCCGAGCCGGGG 481

QY 1466 GCTGAGATGTAGAGCTCGAGGCGGGGGCTCCCGAATCTTCAAAACGACAGACCTGCTCTT 1525

Db 482 GCTGAGATGTAGAGCTCGAGGCGGGGGCTCCCGAATCTTCAAAACGACAGACCTGCTCATAC 541

QY 1526 CGGGGCCCCAGGGGTGTGGGGCGCCGGGGAGACAGCC 1561

Db 542 TCGGGGCCCCAGGGGTGTGGGGCGCCGGGAAACAGCAC 577

RESULT 12	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS
B0086295	B0086295	120h01.1 Melton Normalized Human Islet 4 N4-HIS 1 Home sapiens cDNA clone IMAGE:6135408 5' similar to TR-060843 O60843 PUTATIVE SERINE/THREONINE PROTEIN KINASE ;, mRNA sequence.	B0086295	B0086295.1	GI:20045499
		510 bp	mrna	linear	EST 29-APR-2002

ORGANISM	Homo sapiens
SOURCE	human.
REFERENCE	DOI:

**REFERENCE**  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 510)  
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

TITLE	Endocrine Pancreas Consortium
JOURNAL	unpublished (2000)
COMMENT	Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel.: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@ionp.harvard.edu  
 Library was constructed by Dr. Douglas Melton DNA sequencing by:  
 Washington University Genome Sequencing Center This clone is  
 available royalty-free through LIND; please contact the IMAGE  
 consortium (info@image.llnl.gov) for further information  
 Seq primer: -40RP from Gidco  
 High quality sequence stop: 445.  
 location/Qualifiers  
 1..510

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6135408"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"

```



DEFINITION UI-M-EMO-bx1-p-15-0-UI.r1 NIH\_BMAP\_EM0 Mus musculus cDNA clone  
IMAGE:5708798 5', mRNA sequence.  
ACCESSION B0444032  
VERSION B0444032.1 GI:21247144  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 600)  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:  
http://image.lnl.gov  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

## FEATURES

Seq primer: pyx-5.  
Location/Qualifiers

1..600  
/organism="Mus musculus"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5708798"  
/clone\_1ib="NIH\_BMAP\_EM0"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 15.5 dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/note="Organ: brain; Vector: pYX-Asc; Site:1: EcoR I; Site:2: Not I. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGGCTGGGA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemlin Chlin, Ph.D., program coordinator."

BASE COUNT 145 a 156 c 173 g 125 t 1 others  
ORIGIN

Query Match 16.4%; Score 490.2; DB 14; Length 600;

Best Local Similarity 91.4%; Pred. No. 6.2e-58; Mismatches 49; Indels 2; Gaps 2;

Matches 341; Conservative 0; Mismatches 49; Indels 2; Gaps 2;

359 CAGTCTCTCAAGCTCCAGCAGCTCTACAGCAAGAAATATTGTACTGCTGAG 418  
11 CAGGTGTCAAGCTCCAGCAGCTCTACAGCAAGAAATATTGTACTGCTGAG 70  
419 CAGCTCTGGGGGGTGAAGTATTGACTGCTGTAAGAGGAGAGACTGACCCGAG 478  
71 CAGCTTCTGCTGGTGAAGCTGCTGCTGCTGTAAGAGGAGAGACTGACCCGAG 130  
479 GAGGCCGGAAGTCTCTCCGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 538  
131 GAGGCCGGAAGTCTCTCCGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 190  
539 ATCTGCACAGAGACCTTAAGCCGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTG 598  
191 ATCTGCACAGAGACCTTAAGCCGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTG 250

QY 599 ATTGCAGACTTGGGATGGGCTCCCTGAGTGGGGGACAGCCCTCGAGAGCCAGCTGC 658  
DB 251 ATTCGACACTTGTGTATGGGCTCCGACAGTGGGGGACAGCCCTCGAGAGCCAGCTGC 310  
QY 659 GGGTCCCGCCATTATGCTGTCTGAGAGTGAATTAAGGGGAAAAATATGATGGCCGCG 718  
DB 311 GGGTCCCGCCATTATGCTGTCTGAGAGTGAATTAAGGGGAAAAATATGATGGCCGCG 370  
QY 719 GCAGACATGTGAGCTGTGAGATCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 778  
DB 371 GCAGACATGTGAGCTGTGAGATCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 430  
QY 779 GATGACAGACCTCCGCGCATGCTGCTGAGAGAGTGAAGGGGCGCTCTCCACATGCC 838  
DB 431 GATGACAGAC-ACCTGCGCGCATGCTGAGAGAGTGAAGGGGCGCTCTCCACATGCC 489  
QY 839 CACTTATTCCTCCAGATTGCTCCAGAGCTCCTGAGGGAATGATGAGACCCGCA 898  
DB 490 CACTTATTCCTCCAGATTGCTCCAGAGCTCCTGAGGGAATGATGAGAGCCGCGAG 549  
QY 899 AAAAGCTAGTGTGAGCAAAATTCAGAAACATCCTGTAACCTAGAGCGGGA 950  
DB 550 AAAAGCTAGTGTGAGCA-CAATTCAGAAACATCCTGTAACCTAGAGCGGGA 600

## RESULT 15

BF796285

LOCUS BF796285 753 bp mRNA linear EST 12-JAN-2001  
DEFINITION 602258590F1 NIH\_MGC\_85 Homo sapiens cDNA clone IMAGE:4341838 5',  
mRNA sequence.

ACCESSION BF796285  
VERSION BF796285.1 GI:12101339  
KEYWORDS  
SOURCE  
ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 753)  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Louis Staudt, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:  
http://image.lnl.gov  
Plate: LHM9956 row: c column: 23  
High quality sequence stop: 691.

FEATURES  
source  
Location/Qualifiers

1..753  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4341838"  
/clone\_1ib="NIH\_MGC\_85"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lymph; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC library."

BASE COUNT 156 a 212 c 262 g 123 t  
ORIGIN

Query Match 16.4%; Score 487.8; DB 12; Length 753;

Best Local Similarity 96.8%; Pred. No. 1.1e-57; Mismatches 572; Conservative 0; Mismatches 12; Indels 7; Gaps 7;

QY 943 AGCGGGAAACAGAGCCAGACCCGCTGAGCCAGCCCTGCGCGGATGACAT 1002

Db 167 AGCGGGGAAACAGAGACACCGCTGCTGGAGCCAGCCCTGCGCCCGGGTAGCCAT 226  
OY 1003 GCGAGACCTCCATCCAAAGAGAGTGAACCCGAGCTCCTAGAGATGATGACT 1062  
Db 227 GCGAGACCTCCATCCAAAGAGAGTGAACCCGAGCTCCTAGAGATGATGACT 286  
OY 1063 GGGCTGCTTCAGGAGACCGCAGAGCTGCATCCGAGCTGGCAGTAGAGAGAACCA 1122  
Db 287 GGGCTGCTTCAGGAGACCGCAGAGCTGCATCCGAGCTGGCAGTAGAGAGAACCA 346  
OY 1123 AGAAAGATGATATATATCTGCTTTGGATCGAGAGAGCGGTATCCACACTGTAGAGA 1182  
Db 347 AGAAAGATGATATATATCTGCTTTGGATCGAGAGAGCGGTATCCACACTGTAGAGA 406  
OY 1183 CCAGGACCTGCTCCCGGATGATGTTGACCCCGGAGAGCGTGTGATTCGCCAT 1242  
Db 407 CCAGGACCTGCTCCCGGATGATGTTGACCCCGGAGAGCGTGTGATTCGCCAT 466  
OY 1243 GCTGAGCCCTCAGCGGAGCGGACGACGAGCGGAGTCCATGGAAGTCTGAGCATCAC 1302  
Db 467 GCTGAGCTGTCAGCGGAGCGGACGACGAGCGGAGTCCATGGAAGTCTGAGCATCAC 525  
OY 1303 CGATGCGCGGGGTGGTGGCTCCCTGTACCCACCGGAGCGGCTTTGGAGATGCCACGA 1362  
Db 526 CGATGCGCGGGGTGGTGGCTCCCTGTACCCACCGGAGCGGCTTTGGAGATGCCACGA 585  
OY 1363 CAGCCAGAGATCCGTAAGCTCAGT-GGAGCTTCAGGAGTGTGCTCCAGCCCTTAA 1421  
Db 586 CAGCCAGAGATCCGTAAGCTCAGTGGAGAGCTCCCGGCTGTCTCCAG-CTCTAA 644  
OY 1422 GCAAGCCAAAGAGTCCGGTCTTTTCACCGAGAGCGGGGCTGGAGATGAGGCTC 1481  
Db 645 GCAAG-CAAAGAGTCCGGTCTTTTCACCGAGAGCG-GGGGCTGGAGATGAGGCTC 702  
OY 1482 GAGGGGGGGGCTCCCGACTTCCAAAGCAGAGAGCTGCTTCGGGGCC 1532  
Db 703 GA-GCGGGGGCTCCCGA-TTCCAAAGCAGAGAGCTGCTTCGGGGCC 751

Search completed: April 22, 2003, 00:23:38  
Job time : 2731.75 secs

GenCore version 5.1.4\_P5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 20:17:06 ; Search time 54.0237 Seconds  
(Without alignments)  
13249.433 Million cell updates/sec

Title: US-10-003-690-3  
Perfect score: 2334  
Sequence: 1 atgtcgtccggggcgaagga.....ccaacgggaccctctgccc 2334

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Capext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCrus.COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	966.6	41.4	2908	US-09-930-181-1	Sequence 1, Appl1
2	951	40.7	3364	US-09-930-181-3	Sequence 3, Appl1
3	206.6	8.9	1647	US-09-101-146-44	Sequence 44, Appl1
4	200.2	8.6	1929	US-09-359-161-4	Sequence 4, Appl1
5	198.6	8.5	1742	US-08-557-006C-38	Sequence 38, Appl1
6	198.6	8.5	2652	US-08-557-006C-39	Sequence 39, Appl1
7	198.6	8.5	2761	US-08-557-006C-24	Sequence 24, Appl1
8	182.6	7.8	1736	US-08-557-006C-37	Sequence 37, Appl1
9	182.6	7.8	1783	US-08-557-006C-36	Sequence 36, Appl1
10	177.8	7.6	2698	US-08-677-298-1	Sequence 1, Appl1
11	175.2	7.5	1747	US-08-557-006C-44	Sequence 44, Appl1
12	145	6.2	213	US-09-930-181-5	Sequence 44, Appl1
13	135.6	5.8	2132	US-09-159-385-3	Sequence 3, Appl1
14	135.6	5.8	2132	US-09-166-277-3	Sequence 3, Appl1
15	130.6	5.6	1425	US-09-159-385-4	Sequence 4, Appl1
16	130.6	5.6	1429	US-09-186-277-4	Sequence 4, Appl1
17	116.6	5.0	1547	US-09-523-849-4	Sequence 4, Appl1
18	116.6	5.0	1544	US-09-734-673-1	Sequence 1, Appl1
19	116.6	5.0	2060	US-09-523-849-1	Sequence 1, Appl1
20	107.2	4.6	1599	US-09-256-465-1	Sequence 1, Appl1
21	107.2	4.6	1599	US-09-167-322-3	Sequence 1, Appl1
22	107	4.6	2610	US-09-212-771-1	Sequence 3, Appl1
23	107	4.6	2610	US-09-091-058-1	Sequence 3, Appl1
24	103	4.4	2549	US-09-467-082-3	Sequence 1, Appl1
25	102.4	4.4	1302	US-08-913-050A-2	Sequence 2, Appl1
26	98.2	4.2	1244	US-08-755-728-1	Sequence 1, Appl1
27	98.2	4.2	1244	US-08-974-655-1	Sequence 1, Appl1

28	98.2	4.2	1244	US-09-283-011-1	Sequence 1, Appl1
29	98	4.2	4162	US-08-459-448A-26	Sequence 26, Appl1
30	98	4.2	4162	US-08-459-595A-26	Sequence 26, Appl1
31	98	4.2	4162	US-08-459-504B-26	Sequence 26, Appl1
32	98	4.2	4162	US-08-459-444-26	Sequence 26, Appl1
33	98	4.2	4162	US-09-547-422-26	Sequence 26, Appl1
34	98	4.2	4165	US-07-951-715A-26	Sequence 26, Appl1
35	97.2	4.2	2104	US-09-313-930-1	Sequence 1, Appl1
36	97	4.2	2061	US-09-800-960-1	Sequence 1, Appl1
37	96.8	4.1	4935	US-08-631-097-3	Sequence 3, Appl1
38	96.8	4.1	5886	US-08-810-712-9	Sequence 9, Appl1
39	96.2	4.1	1282	US-08-878-989-12	Sequence 12, Appl1
40	96.2	4.1	1282	US-09-272-796-12	Sequence 12, Appl1
41	95	4.1	3972	US-09-770-170-1	Sequence 1, Appl1
42	94.2	4.0	1456	US-08-749-902-2	Sequence 2, Appl1
43	93.8	4.0	5228	US-09-428-711A-15	Sequence 15, Appl1
44	93	4.0	1962	US-08-924-183-4	Sequence 4, Appl1
45	93	4.0	1962	US-09-488-364-4	Sequence 4, Appl1

#### ALIGNMENTS

```

RESULT 1
US-09-930-181-1
; Sequence 1, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 160 101 VI
; CURRENT APPLICATION NUMBER: US/09/930.181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(2112)
US-09-930-181-1

Query Match      41.4%; Score 966.6; DB 4; Length 2908;
Best Local Similarity 70.1%; Pred. No. 3.2e-212;
Matches 1485; Conservative 0; Mismatches 439; Indels 195; Gaps 6;

QY 74 CCCAGCAGCCCAATATATGTGGGCCCTATCGCTGAGAGACGCTGGGCAAGAGACAGA 133
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 134 CGCAGCAGCGCCAGTATGTTGGGCCCTACCGGCTGGAAGACGCTGGGCAAGGCGACA 193

QY 134 CAGGCGTGGTAAATCGGGGGTCCATCGATCAGCGGTGAGAGAGTGGCCATCAAGATCG 193
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 194 CAGGCTGGTGAACCTGGGGGTTCACTCGTACCTGCAGAGAGTGGCCATCAAGATCG 253

QY 194 TGAACCGGAGAACCTGTGCGAGTGTGCTGATGAAGTGGAGCGGAGATCGCATCC 253
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 254 TCAACCGTGAAGAACCTGACGAGTGTGCTGATGAAGTGGAGCGGAGATCGCATCC 313

QY 254 TGAAGCTGATGAAACCCACCATGTCTCAAGCTCCAGAGCTTACGAGAAAGAAAT 313
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 314 TGAAGCTGATGAAACCCACCATGTCTCAAGCTCCAGAGCTTACGAGAAAGAAAT 373

QY 314 ATTGTGACCTGTTCTGGAGCAGCGTGGGGGGTGGAGCTATTCAGTACCGGTAAAGA 373
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 374 ATTGTGACCTGTTCTGGAGCAGCGTGGGGGGTGGAGCTATTCAGTACCGGTAAAGA 433

QY 374 AGGGAAGCTGACCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 433
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 434 AGGGAAGCTGACCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 493

QY 434 ACTTCTGCACAGCTACTCTCATCTGCGACAGAGACCTTAAAGCCGAGAACTGCTTTGG 493

```

```

Db 494 ACCTTCCACAGGACCTCCATATGCCAGAGGATCTGAAAACCTGCTCTGG 553
OY 494 ATGAGAAAAAACAATCCGATTCGACACTTGGAGTGGGCGCTCGAGGGGAGCA 553
Db 554 ACAGAGAAACACATCCGATCCGACACTTGGATGGCGCTCGAGGGTGGAGCA 613
OY 554 GCCTCTGAGACAGCTGGGGTCCCGCACTTATGCTGTCCAGAGGTGATTAGGGG 613
Db 614 GCCTGTTGAGACAGCTGTGGTCCCGCACTTACGCTGCGCCGAGGATCCGGGGG 673
OY 614 AAAATATATGCGCGCGCGGACAGCATGTGAGCTGTGAGTACTCTTCCGCTGC 673
Db 674 AGAGATATGACGGCGGAGAGGCGAGCTGTGAGCTGTGAGTACTCTTCCGCTGC 733
OY 674 TCGTGGGCGCTGCTCCCTTGTATGACGACACATCCCGCACTGTGTGAGAGTGAAC 733
Db 734 TGTGTGGGGGCTCTGCTCTTGCAGATGACATCTTGCAGAGCTGTGAGAGGCA 793
OY 734 GGGGCGCTCTTCCACATCCCGCACTTCTCTCAGATTTGCCAGAGCTCTGAGGGGA 793
Db 794 GGGGCGCTCTTCCACATCCCGCACTTCTCTCAGATTTGCCAGAGCTGTGAGAGGCA 853
OY 794 TGTATGAGTGGAGCGCGGAAAAAGGCTCACTGTGAGCAATTTAGAAATCTTGTGT 853
Db 854 TGTATGAGTGGAGCGCGGAAAAAGGCTCACTGTGAGCAATTTAGAAATCTTGTGT 913
OY 854 ACCTAGGCGGAAAAACAGAGACAGACCCGCTGTGAGCAAGCCCTGCGCGCGGGTAG 913
Db 914 ATATAGGGGGGAGAAATGAGAGCCGAAAC-----AGAGCAGCCCTTCTCTGCAAGGTGC 967
OY 914 CCATGCGAGCTGCTCATCCAGAGAGTGTGAGCCGAGCTCTCTAGAGAGCATGGCAT 973
Db 968 AGATCCCTGCTGCTGAGCCTGTGAGAGCATGAGACCCGAGCTGTGAGAGCATGCACT 1027
OY 974 CACTGGGCTCTTCAAGGACCGCGAGAGGCTGTGATCGGAGCTGTGAGAGGAGAGA 1033
Db 1028 CACTGGGCTCTTCAAGGACCGCGAGAGGCTGTGATCGGAGGAGGAGAGA 1087
OY 1034 ACCAAGAAAGATGATATATATCTGCTTGTGATCGAGAGAGGAGGATATCCACCTGTG 1093
Db 1088 ACCAGAGAAATGATATATCTGCTTGTGATCGAGAGGAGGAGGAGGAGG 1147
OY 1094 AGACACAGAGCTGCTCCCGGAGATGTTGAGCCCGCGGAGAGGCTGTGATTTCTC 1153
Db 1148 AGATGAGAGCTGCTCCCGGAGATGTTGAGCCCGCGGAGAGGCTGTGATTTCTC 1207
OY 1154 CCATGCTGAGCTGCTCATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1213
Db 1208 CGATGCTGAAACCGGACGCGAGGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGG 1267
OY 1214 TCACGATGCGCGGGGCTGCTCCCGGAGATCCGAGCCGAGGAGGAGGAGGAGGAGG 1273
Db 1268 TGAC-----GAGCGGGGCTGCTCCCGGAGATCCGAGCCGAGGAGGAGGAGGAGG 1318
OY 1274 AGCAGACGCGAGATCCGATGAGCTGTGAGAGGCTCTCCAGGAGGCTGTCTCCAGCCCTC 1333
Db 1319 AGCAGACGCGAGATCCGATGAGCTGTGAGAGGCTCTCCAGGAGGCTGTCTCCAGCCCTC 1378
OY 1334 TAGAGAGCCCAAGAGTCCGCTTCTTCTTTCACCGGAGCGGGGCTGTGAGATGAGG 1393
Db 1379 TCAGAGAGCCCG----- 1391
OY 1394 CTCGAGAGGCGGCTCCCGCACTTCAAAAGCAGAGCGCTCTTCTGAGGCGCCAGAGG 1453
Db 1392 ----- 1391
OY 1454 GTGCGGCGCGCGGAGAGCCCGCGCGCGCGCTCCAGAGGCGCGCTCCAGAGCCCTGCGCGGC 1513
Db 1392 -----GCTGACCCCTTCAACCCCTTCAACCAAGGAGGAGTCCCTTCCCG----- 1431
OY 1514 CCCAGAGCTCCCGCGCTCTCTGTGGGCGGAGCCCGCTTGCACCTGTGAGACAGCGCCC 1573

```

```

Db 1432 -----ACCCCAAGGGGACACTGTCCACAGCCAA 1462
OY 1574 GGGCAGTCCACCGGGAGCCCGGGAGCAACACACCCCGCGCGGTGGCTGC 1633
Db 1463 AGAGAGACCCGCTGGCAGCCCAACCCAGCCCGCGTCCAGCC-----AGCGTCG 1516
OY 1634 GGGAGCCCGCTGGAGAGTGTCTCAACTCCATCCGACAGGCTTCTGGGCTCCCTC 1693
Db 1517 GAGGGGTGCCCTGGAGGGCGCGGCTCAACTCCATCCAGAGGAGGAGGAGGAGGAGG 1576
OY 1694 GCTTTACCGCGCGCAGATCAGTCCCTTACCGCTGTGAGAGATGTCCAGCTTGAACCCAG 1733
Db 1577 GCTTTACCGCGCGCAGATCAGTCCCTTACCGCTGTGAGAGATGTCCAGCTTGAACCCAG 1636
OY 1754 AGTCCTCCCGGAGCTGGGCAAAAGCGCTCGGTTCCGGAAGCTTATCTCTTGAACAAG 1813
Db 1637 AGTCCTCCCGGAGCTGGGCAAAAGCGCTCGGTTCCGGAAGCTTATCTCTTGAACAAG 1636
OY 1814 AAGAAATATTTCTCTGCTTAAAGCAAACTCTCAGAGCATCAAAAGCAGATCG 1873
Db 1697 AGAGAGAGATCTTGTGTGATCATAAAGCAAACTCTGAGCTCATAGGCTGACATCG 1756
OY 1874 TCCATGCTTCTGTGTGATCCCAAGCTGTGATCAGTGTGCTGTCAAGACAGCTTCA 1933
Db 1757 TGCAGCGCTTCTGTGTGATCCCAAGCTGTGATCAGTGTGCTGTCAAGACAGCTTCC 1816
OY 1934 GGGCGAGTACAAAGGCGAGTGGGCGCGCTCGCTTCAAAAGCGCTTCCAGG 1993
Db 1817 GGGCGAGTACAAAGGCGAGTGGGCGCGCTCGCTTCAAAAGCGCTTCCAGG 1876
OY 1994 TGGACATCAGCTCTGTGAGGCTCCAGAGCCCTCCCGGAGGAGGAGGAGGAGGAGG 2053
Db 1877 TTGATATCAGCTTACAGGAGGCT-----GGGAGAGCGGAGAGAGGA 1918
OY 2054 GTGGCATCTACTCGCTGACCTTCACTCTCATCTGCGGTCCAGCGCTTCAAGGAG 2113
Db 1919 ACGGCATCTACTCGCTGACCTTCACTCTCATCTGCGGTCCAGCGCTTCAAGGAGG 1978
OY 2114 TGGTGGAGACATCCAGGACAGCTCTGTGAGCACTCATGTGACCAAGCCCTCGTGAAGGCC 2173
Db 1979 TGGTGGAGACATCCAGGACAGCTCTGTGAGCACTCATGTGACCAAGCCCTCGTGAAGGCC 2038
OY 2174 TGGAGAGCGAGAAAGG 2192
Db 2039 TGTGAGAGCACTAAGT 2057

RESULT 2
US-09-930-181-3
; Sequence 3, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Orligene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase In Brain and Pancreas
; FILE REFERENCE: 16U 101 VI
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (482)..(2239)
US-09-930-181-3

Query Match
Best Local Similarity 69.9%; Score 951; DB 4; Length 3364;
Matches 1484; Conservative 0; Mismatches 440; Indels 199; Gaps 7;
OY 74 CCCAGCAGCCCAATATGTGGGCGCCCTATGCTGTGAGAAAGAGCTGGCAAGAGACGA 133

```

Db 257 CGCGACGGGCGATATGTTGGGCCCTACCGGTGAGAGAGCGTGGGCAAGGGGCA 316  
 QY 134 CAGGCGTGTAACTGGGGTCCATCGATCAGGGGTGAGAGGTCCCATCAAGATG 193  
 Db 317 CAGGTCTGGTGAAGCTGGGGTTCACCTCCGACCTGCGAGAGGTGGCCATCAAGATG 376  
 QY 194 TGAACCGGGAAGCTGTGGAGTGGTGTGATGAGGTGAGGCGGAGATGGCATCC 253  
 Db 377 TCAACGCTGAGAGCTCAGGAGTGGGTGTGATGAGGTGAGGCGGAGATGGCATCC 436  
 QY 254 TGAAGCTCATGAGACCCCATATGCTTCAAGCTCCAGAGGTCTACAGAGAAACAAGAT 313  
 Db 437 TGAAGCTCATGAGACCCCATATGCTTCAAGCTCCAGAGGTCTACAGAGAAACAAGAT 496  
 QY 314 ATTT-----GTACCTGGTTCGTGAGACGCTTCGGGGGGTGAAGCTATTCGACTACCTGTA 369  
 Db 497 ATTTGTAGTACCTGGTTCGTGAGACGCTTCGAGTGGTGAAGCTTCGACTACCTGTA 556  
 QY 370 AAGAAGGGGAGACTGACGCCCAAGAGGCGCGAAAGTTCTTCGCGAGATTTGTCTGG 429  
 Db 557 AAGAAGGGGAGGCTGAGCGCTTAAGAGGCTGGAAGTTCTTCGCGAGATTCATCTGCG 616  
 QY 430 CTGGACTTCGCGACGCTACTCATGTCGACAGAGACCTTAAAGCCGAGAACCTGCTT 489  
 Db 617 CTGGACTTCGCGACGCTACTCATGTCGACAGAGACCTTAAAGCCGAGAACCTGCTT 676  
 QY 490 TTGAGTGAAGAAAAACAACATCCGATTCGAGACTTCGCGATGGCTCCCTGCAAGTGGG 549  
 Db 677 CTGGAGAGAAAGAACAAACATCCGATTCGAGACTTCGCGATGGCTCCCTGCAAGTGGG 736  
 QY 550 GACAGCTCTGAGACCAAGCTGCGGGTCCCGCATTAAGCTGTTCAGAGGTGATTAG 609  
 Db 737 GACAGCTCTGAGACCAAGCTGCGGGTCCCGCATTAAGCTGTTCAGAGGTGATTAG 796  
 QY 610 GGGGAAAAATATGATGGCGCGCGGAGACATGTTGAGTGTGAGTCACTCTCTGGCC 669  
 Db 797 GGGGAAAAATATGATGGCGCGCGGAGACATGTTGAGTGTGAGTCACTCTCTGGCC 856  
 QY 670 CTGCTCTGGGGGCTTCGCTTTGATGACACAACCTCCGACAGCTGCTGAGAGATG 729  
 Db 857 TTGCTGTGGGGGCTTCGCTTTGATGACACAACCTCCGACAGCTGCTGAGAGATG 916  
 QY 730 AAACGGGCGCTTCACACAGCCCACTTATCTCTCAAGTTCGACAGCTCTCTGAG 789  
 Db 917 AAAGGGGCGGCTTCACACAGCCCACTTATCTCTCAAGTTCGACAGCTCTCTGAG 976  
 QY 790 GAATGATTCGAGTGGAGCCCGAAAAAGGCTCACTGAGCAATTCAGAAACATCT 849  
 Db 977 GGCATGATTCGAGTGGAGCCCGAAAAAGGCTCACTGAGCAATTCAGAAACATCT 1036  
 QY 850 TGTACTTGGGGGAGGAAACAGAGCCAGACCTGCTGAGACCAAGCTCTGGCGCGG 909  
 Db 1037 TGTACTTGGGGGAGGAAACAGAGCCAGACCTGCTGAGACCAAGCTCTGGCGCGG 1090  
 QY 910 GTAGCATTCGAGGAGCTTCACAGCCGAGAGCTGAGACCCCGAGCTCTGAGAGCATG 969  
 Db 1091 GTAGCATTCGAGGAGCTTCACAGCCGAGAGCTGAGACCCCGAGCTCTGAGAGCATG 1150  
 QY 970 GCATCATTCGAGGAGCTTCACAGCCGAGAGCTGAGACCCCGAGCTCTGAGAGCATG 1029  
 Db 1151 GCATCATTCGAGGAGCTTCACAGCCGAGAGCTGAGACCCCGAGCTCTGAGAGCATG 1210  
 QY 1030 GAGAACCAAGAAAGATGATATATATCTGTTTGAATCGAGAGAGGCTATCCAC 1089  
 Db 1211 GAGAACCAAGAGAGATGATATATCTGTTTGAATCGAGAGAGGCTATCCAC 1270  
 QY 1090 TGTGAGACCAAGAGCTTCCTCCCGAATGATGTTGACCCCGAGAGGCTGTGAT 1149  
 Db 1271 TGTGAGACCAAGAGCTTCCTCCCGAATGATGTTGACCCCGAGAGGCTGTGAT 1330  
 QY 1150 TCTCCCATCTGAGACCCGTCAGGAGGAGGCGAGACCAAGCCGAATCTCATGAGTCTG 1209

Db 1331 TCCCCGATGCTGAACCGGACGGCAAGCGGCGCCAGACGCAAAATTCATGAGTGTCTC 1390  
 QY 1210 AGCATCACCATGCGGGGGGTGTGGCTCCCTGTACCCACCGAGCGGCTTGGAGATG 1269  
 Db 1391 AGGTGAC-----GAGCGGGGCTCCCGGGTGGCTGGCGGGGCGCATTAAGATG 1441  
 QY 1270 GCCCAGCACAGCCAGAGATCCGTAAGCTCAGTGAAGCTCCAGGGTCTGCTCCAGC 1329  
 Db 1442 GCCCAGCACAGCCAGAGATCCGTAAGCTCAGTGAAGCTCCAGGGTCTGCTCCAGC 1501  
 QY 1330 CCTCTAAGACGCCAAGAGAGTCCGCTCTTTCTTTTACCGGAGCCGGGGCTGAGAT 1389  
 Db 1502 CCACTCAGCAGCC----- 1514  
 QY 1390 GAGGCTGAGAGCGGGGCTCCCGACTTCAAAAGCAGACGCTGCTTCGGGGCCCC 1449  
 Db 1515 ----- 1514  
 QY 1450 AGGGGTGGGGGCGCGGGAGACGCCCGCCCGCACTGCGCTCCACACCCCTGCC 1509  
 Db 1515 -----CCGGGTGACCCCTCACCCCTCACAAAGGGGCAAGTCCCTCCCGC----- 1558  
 QY 1510 GGGCCCCCAGGCTCCCGGCTCTCTGGGGGAGCCCGCTGCACTGGCTGTGACAG 1569  
 Db 1559 -----ACCCCAAGGGAGACACTGTCCACAGC 1585  
 QY 1570 CCGCGGCGACGTCGCCACCGGAGCCCGGGAGCAACACCCAGCCCGCGGGGCTGGC 1629  
 Db 1586 CCAAGGAGAGAGCCGGGTGGAGAGCCCAAGCCCAAGCCCGCGGTCCAGGCC-----AGC 1639  
 QY 1630 GTGGGGGAGCGCGCTGAGAGATGCTCTCAACTCATCCGCAAGAGTCTCTGGGCTCC 1689  
 Db 1640 GTGGGGGAGCGCGCTGAGAGAGCGCGGTCAACTCAAGCAAGAGTCTCTGGGCTCA 1699  
 QY 1690 CCTGCTTTCACCGGCGAAGATGAGTCCCTACCGCTGGAGAGATGCCAGTGTAGC 1749  
 Db 1700 CCGCGCTTCACCGCGCGGAAACTGCAAGTTCGAGCGCGGAGAGATGTCACACTGACA 1759  
 QY 1750 CCAAGTCTCCCGCGGAGCTGGCAAAAGCTCTGTGGTTCGGGAATTCATCTCTTGGAC 1809  
 Db 1760 CCAAGTCTCCCGCGGAGCTGGCAAAAGCTCTGTGGTTCGGGAATTCATCTCTTGGAC 1819  
 QY 1810 AAAGAGAACAAATATTCCTCTGCTTAAGAGCAAAACCTCTCAGCAGATCAAGCAGAC 1869  
 Db 1820 AAAGAGAGAGAGATCTCTGCTGATCAAAAGCAAAACCTCTGAGCTCATCAAGCTGAC 1879  
 QY 1870 ATGCTCATGCTCTTGTGATCCCGAGCTGAGTCAAGTGTCTGACAGACAGC 1929  
 Db 1880 ATGCTCATGCTCTTGTGATCCCGAGCTGAGTCAAGTGTCTGACAGACAGC 1939  
 QY 1930 TTCAGGCGCGAGTACAAGGCGAGTGGCGGCCCTCCGCTTCCAAAGCCCTCCGCTTC 1989  
 Db 1940 TTCAGGCGCGAGTACAAGGCGAGTGGCGGCCCTCCGCTTCCAAAGCCCTCCGCTTC 1999  
 QY 1990 CAGGTGACATCAGCTCTCTGAGGCTCAGAGCCCTCCCGCAGCGAGCGGACGGA 2049  
 Db 2000 CAGGTGATATCAGCTCAGCAGAGGT-----GGGGAGCGCGAGAG 2041  
 QY 2050 GGTGGTGCATCTACCTCGTACCTTCACTCATCTCGGGGCCAGCGGTGGTTCAG 2109  
 Db 2042 GAGAACGGCATCTACCTCGTACCTTCACTCATCTCGGGGCCAGCGGTGGTTCAG 2101  
 QY 2110 CGAGTGTGAGACATTCAGGACAGAGTCTCTGAGACTTGAACAGCCCTCCGTGAC 2169  
 Db 2102 AGGGTGTGAGACATTCAGGACAGAGTCTCTGAGACTTGAACAGCCCTCCGTGAC 2161  
 QY 2170 GCGCTGACAGCAAGAACG 2192  
 Db 2162 CACTTGTACAGACCACTACTG 2184

RESULT 3  
 US-09-101-146-44

```

; Sequence 44, Application US/09101146
; Patent No. 6124125
; GENERAL INFORMATION:
; APPLICANT: Dartmouth College, St. Vincent's Institute of
; APPLICANT: Medical Research, Kemp et al.
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 66 E. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PC
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/101,146
; FILING DATE: October 7, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PNT/450
; FILING DATE: 8 JAN 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: DC-0050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (856) 810-1515
; TELEFAX: (856) 810-1454
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1647
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: NO
; US-09-101-146-44

Query Match      8.9%; Score 206.6; DB 3; Length 1647;
Best Local Similarity 55.3%; Pred. No. 2,4e-38;
Matches 423; Conservative 0; Mismatches 339; Indels 3; Gaps 1;

QY 92 TGGGCCCTATCGGCTGAGAAAGCGCTGGGCAAGAGACAGACAGCGCTGTTAAACTCG 151
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 38 TGGGCCCTATCGGCTGAGAAAGCGCTGGGCAAGAGACAGACAGCGCTGTTAAACTCG 97
QY 152 GGGTCCATGTCATCAGCGGTGAGAGAGTGGCCATCAAGATCGTGAACCGGAGAACTGT 211
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 98 GCAAGCAGCAGTGTGAGTGAACATGCTGCTGGAAGATCTCAACCGGAGAAATTC 157
QY 212 CGGAGTCGTCGAT---GAAAGTGAAGCGGAGATCGCATCTGTAAGTCAATCGAAC 268
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 158 GAAAGCTGGAGCGGTGCGGAAATCCGAGAGATCCAGAACCTGGAAGCTTTTCAGGC 217
QY 269 ACCCATGTCCTCAAGCTCCAGACGCTTACAGAGAACAAAGAAATATTTGACCTGTTG 328
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 218 ACCCTATATATCAAACTTACAGAGTCAATACAGCTTGTATTTTCAATGTCAC 277
QY 329 TGGAGCAGCTCTCGGGGGGTGAGAGTATTCAGCTGTAAGAAAGGAGAGTACGCG 388
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 278 TGGAAATATGTGAGAGAGAGTATTTGATCTGTAAGAAAGGAGTTCGAGC 337
QY 389 CCAGAGAGCGCCAAAGTCTTCCGCAAGATGTGTCTGCGCTGAGCTTCCACACT 448
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 338 AAAAGAGAGTCAAGCTCTCTCCAGAGATCTTTCTGCTGATGTCACAGAGC 397
QY 449 ACTCATCTCCACAGAGACTAAAGCCGAGAACTGCTTTGGATGAGAAAAACAACA 508
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 398 ATATGTGTCTCACAGAGATTTGAACCTGAAACGCTCTCTGATGACACATGAAATG 457

```

```

QY 509 TCCGATTCAGACTTCGCGATGCGCTCCCTGCAGGTGGGAGACGCTCTGGAGACA 568
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 458 CAAGATATGCGGCACTTGGCTTTCAAAACATGATGATGATGATTTTAAAGACA 517
QY 569 GCTGGGGTCCCCCATTTATGCTGTCCAGAGTGAATTAAGGGGAAAAATATGATGCGC 628
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 518 GCTGTGCTGCGCCCAATTTATGTCACACAGAAATTAATTCAGAAATTTCAAGACGCC 577
QY 629 GCGGCGAGCATGATGAGTGTGAGTCAATCCCTTCGCGCTCTGCTGGGCGCTCGC 688
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 578 CTGAGTATGACATCTGGAGAGCGGCTCATCTCTATGCTTGTGAGAACTCTCC 637
QY 689 CTTTATGATGACAAACCTCCGCGAGCTGTGAGAGATGAACGGGGCTCTCCACA 748
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 638 CTTTATGATGATGACCAAGCTGCTTTTAAAGATATGACAGGGAATTTTATA 697
QY 749 TGGCCCACTTCATCTCCAGATTGCCAGAGCTCTGTAGGGGATATCGAAGTGAAGC 808
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 698 CCGCTCAGATTTGAATCCCTGTATATAAGCCTTTGAAACATATCTGAGGTAGTC 757
QY 809 CCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTCTGTGT 853
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 758 CTATGAGAGGCGCCACATTAAGATATCAGGAGACATGATGAT 802

RESULT 4
US-09-359-161-4
; Sequence 4, Application US/09359161A
; Patent No. 6342656
; GENERAL INFORMATION:
; APPLICANT: Bradford, Kent J.
; APPLICANT: Dahal, Peetambar
; APPLICANT: Yang, Hong
; APPLICANT: Cooley, Michael
; APPLICANT: Downie, Bruce
; APPLICANT: Gee, Oliver
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses
; FILE REFERENCE: 023070-095900US
; CURRENT APPLICATION NUMBER: US/09/359,161A
; CURRENT FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1929
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; OTHER INFORMATION: SNF1 kinase subunit of protein kinase (LesNf1)
; US-09-359-161-4

Query Match      8.6%; Score 200.2; DB 4; Length 1929;
Best Local Similarity 54.4%; Pred. No. 7,5e-37;
Matches 425; Conservative 0; Mismatches 353; Indels 3; Gaps 1;

QY 82 GCCCAATATGTGGGCCCTATCGCTGAGAGACGCTGGGCAAGAGACAGAGGCTG 141
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 69 GACTCATTTTTCAGGACATTAATCACTCGGAAACACTTGGATGTGATGCTTGGCAAA 128
QY 142 GTTAAATCGGGGCTCAGTGCATCAGGGGTGAGAGGTCCCATCAAGATCTGAACCG 201
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 129 GTTAAATATGCTGAACATACGTTTAACAGGCGACAAAGTTCTGTCAAGATTTCTTAATCGT 188
QY 202 GAGAACTGTGAGTGG---TGTGATGAGGTGAGCGGAGATCGCCATCTGAAG 258
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 189 CGAAAAATCAGGAATATGACATGAGAGAAAGTCCGTAGAGAAATCAAAATATTTGAGA 248
QY 259 CTTATGAAACCCATGATGCTCAAGCTCCAGAGCTTACAGAGAAATATTTG 318
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 249 TTGTTATCATCTCTATTTATATGAGCTTTATGAGGTCTATAGAGACATCATGATATA 308

```

QY 319 TACCTGCTTGTGAGCAGCTCTCGGGGCTGACCTATTGCACTACCTGTAAAGAGGG 378  
 DB 309 TATGTTGTGTGAGTGTGAAATCTGGCGAGTATTGATTGATTGAGAGGGC 368  
 QY 379 AGACTGACGCCCAAGAGGCGGAGGCTTTCGGCGAGTTGTGTGCGCTGTGACTTC 438  
 DB 369 AGATTGACAGAGATGAGTACCTGTAATTTTTCAGCAGATATTTCTGTGTGAGTAC 428  
 QY 439 TGGCAGAGCTACTCTCATCTGCCACAGAGACCTAAAGCCGGAACCTCTTTTGTGATGAG 498  
 DB 429 TGGCATGAGAAACATGTGTGTATAGAGACTTAAGCTGAAACCTCTCTGTGTGACTCC 488  
 QY 499 AAAAAACATCTCCGATTTGCACTGTGGCATGCGGCTCCCTGCAAGGTGGGAGACGCTC 558  
 DB 489 AATGTGATGTGAGATGCGAGATTTGTTGATGACATATCATGCGGAGTGTCTATTT 548  
 QY 559 CTGAGAGCAGCTCGGGGCTGCTCCATTTATGCTGTCAGAGGTGATTAAGGGGAAAAA 618  
 DB 549 CTGAGAGCAGTGTGCGGAGAGCCCAACTATGCTGCGCCAGAGGTATATCAGGTAAATTG 608  
 QY 619 TATGATGCGCGCGGAGACATGTGAGCTGTGAGTCACTCTTCCGCTGCTGTG 678  
 DB 609 TATGCTGCGCTGAGTATGATGAGTGTGTGTTATCTTATGCTCTCTGTG 668  
 QY 679 GGGGCTCTGCTTGTGATGAGACACCTCCGCACTGCTGTGAGAGAGTAAAGGGGC 738  
 DB 669 GGGACCTTCCGTTGAGTATGAGTAAACATACCAATCTTTTAAAGAAATTAAGGTGTA 728  
 QY 739 GTCCTTCCACATGCCCCCACTTCATCTCCAGATTGCGACAGCCTCTGAGGGAGTATC 798  
 DB 729 ATATATATCTGCTCCAGCATTATTCAGCTGTGTGAGGAGATTGATTCAGAGTACTT 788  
 QY 799 GAAGTGAAGCCGCAAAAAGGCTCAGTGTGAGCAAAATTCAGAAACATCTGTGACTTA 858  
 DB 789 ATATGACACCAATGAAGCGAATGATCTCTGATGATTGCGCTGACCCCTTGTTCCAA 848  
 QY 859 G 859  
 DB 849 G 849  
 RESULT 5  
 US-08-557-006C-38  
 ; Sequence 38, Application US/08557006C  
 ; Patent No. 6258547  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berl, Rajindar K.  
 ; APPLICANT: Carling, David  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
 ; FILE REFERENCE: NGAP/PM37588/UST  
 ; CURRENT APPLICATION NUMBER: US/08/557,006C  
 ; CURRENT FILING DATE: 1996-03-06  
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093  
 ; PRIOR FILING DATE: 1994-05-20  
 ; PRIOR APPLICATION NUMBER: GB 9310489.1  
 ; PRIOR FILING DATE: 1993-05-21  
 ; PRIOR APPLICATION NUMBER: GB 9318010.7  
 ; PRIOR FILING DATE: 1993-08-31  
 ; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: Patentln Ver. 2.1  
 ; SEQ ID NO 38  
 ; LENGTH: 1742  
 ; TYPE: DNA  
 ; ORGANISM: Human AMP protein kinase  
 ; US-08-557-006C-38  
 Query Match 8.5%; Score 198.6; DB 4; Length 1742;  
 Best Local Similarity 54.6%; Pred. No. 1.7e-36;  
 Matches 418; Conservative 0; Mismatches 344; Indels 3; Gaps 1;  
 QY 92 TGGCCCTATTCGGCTGAGAGAGCTGGCGCAAGACAGACAGCGCTGTTAAACTCG 151

DB 38 TCGGACACTAGCTGTGGGGAGACACCTGGGCTGCGACCTTGCGCAAGTGAAGATTG 97  
 QY 152 GGGTCCACTGCATCAGCGGTGAGAGGTGCGCATCAAGATCGTGAACCGGGAGAGCTGT 211  
 DB 98 GAGACATCAATTAGACGGGCATTAAGTGGCAGTTAGATCTTAAATGACAGAAAGATTG 157  
 QY 212 CGGAGT---CGGTCTGATGATGAGGTGAGCGGAGATGCCATCCTGAGACTCATGAAAC 268  
 DB 158 GCAGTTAGATGTTGTGAAAAAATAAGAGAAATTCAAAATCTTAACTCTTCCTC 217  
 QY 269 ACCGACATGCCCAAGCTCAGAGCTGACAGCTGATGACACTCCAGAGACTTTTATGTTA 328  
 DB 218 ATCTCATATATCAAACTCTTACCAAGTATGACACTCCAGAGACTTTTATGTTA 277  
 QY 329 TGGAGACGCTCTCGGGGGTGTGATCTATTCGACTCTGTTAAAGAGGAGACTGAGC 388  
 DB 278 TGGAAATATGTCTGAGAGGTGAATTGTTCGATCATCTGTAACACCGGAGGTTGAAG 337  
 QY 389 CCAAGAGCGCCGAAAGTTCTTCGCGCAGATTTGTCTGCGCTGACTTGTGCCACAGCT 448  
 DB 338 AGGTGAGAGCTGCGCGCTCTTCCAGCAGATTCGTCTGCGGACTACTGTACAGGC 397  
 QY 449 ACTCATCTGCGACAGAGACTTAAAGCCGGAAGACCTGTTTGGATGAGAAAAACA 508  
 DB 398 ACATGTTGTTCACAGGAGCTGAGCAGAGAACGTTGTTCTGAGACGCCAGATGATG 457  
 QY 509 TCCGATTGCGAGCTTGGGATGCGCTCCCTGCAAGGTGGGAGAGCCTCTGAGACCA 568  
 DB 458 CTAAAGATAGCTGACTTGGACTCTCTAATATGATGTCAGATGATGATTTCTACAGACTA 517  
 QY 569 GCTGCGGGTCCCCCATTTATGCTGTCCAGAGTGTATTAAGGGGAAAAATATGATGCGC 628  
 DB 518 GCTGTGATGCCAAATTAATGACAGCAGCGAGTCACTCTCGAAGAGCTGTATGGGGTTC 577  
 QY 629 GCGGGCAGATGTGAGGTGTGAGTCAATCCTTCCGCTGCTGTGGGGCTGCTGC 688  
 DB 578 CTGAGTTGATATCTGAGCTGTGTGTATCTGTATGCTCTCTGTGACACCTTC 637  
 QY 689 CTTTGTGAGCAGACACCTCGGAGCTGTGAGAAAGTGAAGCGGGCTCTTCCACA 748  
 DB 638 CTTTGCAGATGAGCAGCTGCTACGCTCTTAAGAAATCCAGGGGGTGTCTTACA 697  
 QY 749 TGGCCCACTTCATCTTCCAGATTTCCAGAGCTCTCTGAGGGGAATGATGAGTGAAGC 808  
 DB 698 TCCCGGAGTATCCACCGTTCTATTCACACTCTGTGATGACATGCTGAGAGTGGAGC 757  
 QY 809 CCGAAAAAGGCTCAGCTGTGAGCAATTCAGAAACATCCTTGT 853  
 DB 758 CTTGAAAGCAGCAACTATCAAGACATAGAGAGCATGAATGTT 802  
 RESULT 6  
 US-08-557-006C-39  
 ; Sequence 39, Application US/08557006C  
 ; Patent No. 6258547  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berl, Rajindar K.  
 ; APPLICANT: Carling, David  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
 ; FILE REFERENCE: NGAP/PM37588/UST  
 ; CURRENT APPLICATION NUMBER: US/08/557,006C  
 ; CURRENT FILING DATE: 1996-03-06  
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093  
 ; PRIOR FILING DATE: 1994-05-20  
 ; PRIOR APPLICATION NUMBER: GB 9310489.1  
 ; PRIOR FILING DATE: 1993-05-21  
 ; PRIOR APPLICATION NUMBER: GB 9318010.7  
 ; PRIOR FILING DATE: 1993-08-31  
 ; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: Patentln Ver. 2.1  
 ; SEQ ID NO 39



Db 721 TCCGGAGATATCTACACCGTCTATATGCCACTCTGCTATGACATCTGCGAGTGGACC 780  
 Oy 809 CCGAAGAAAGGCTCAGCTGGAGCAAAATTCAGAAACATCCTTGT 853  
 Db 781 CCTGAAGCGAGCAACTATCAAGACATACGAGACATGATGTGT 825

## RESULT 8

US-08-557-006C-37  
 ; Sequence 37, Application US/08557006C  
 ; Patent No. 6258547  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berl, Rajindar K.  
 ; APPLICANT: Carling, David  
 ; APPLICANT: Forster, Robert A.  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
 ; FILE REFERENCE: NGAP/PH37588/OST  
 ; CURRENT APPLICATION NUMBER: US/08/557,006C  
 ; CURRENT FILING DATE: 1996-03-06  
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093  
 ; PRIOR FILING DATE: 1994-05-20  
 ; PRIOR APPLICATION NUMBER: GB 9310489.1  
 ; PRIOR FILING DATE: 1993-05-21  
 ; PRIOR APPLICATION NUMBER: GB 9318010.7  
 ; PRIOR FILING DATE: 1993-08-31  
 ; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 37  
 ; LENGTH: 1736  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Rat liver AMP  
 ; OTHER INFORMATION: protein kinase  
 US-08-557-006C-37

Query Match 7.88; Score 182.6; DB 4; Length 1736;  
 Best Local Similarity 53.38; Pred. No. 7.7e-33;  
 Matches 408; Conservative 0; Mismatches 354; Indels 3; Gaps 1;

Oy 92 TGGCCCCCTATCGGCTGAGAGAGAGCTGGGCAAGAGACAGAGGGCTGTTAACTCG 151  
 Db 38 TCGGACACTAGCTGCTGGGCGACAGCTGGGCGTGGCGACCTTCGGCAAGAGAAATTG 97  
 Oy 152 GGGTCACCTGCATCAGCGGTCAGAGAGTCCCATCAAGATCGTGAACCGGAGAACCTGT 211  
 Db 98 GAGAACATCAATTAACAGGCCATTAAGTGGCAGTTAAATCTTAATAGACAGAAATTC 157  
 Oy 212 CGAGT---CGGTGCTGATGAAGGTGAGCGGAGATGCCATCCCTGAAGCTCATCGAAC 268  
 Db 158 GCACTTATAGATGTTGTTGAAAAAATAAAGCAAGAAATTCMAAATCTMAAACTCTTCGTC 217  
 Oy 269 ACCCAGATGCTCTCAAGCTCCACGAGCTACAGAGAACAAATATTTGTAACCTGGTTC 328  
 Db 218 ATCTCATATATTAACAATAACAGAGTATGATGACATCCCAACAGATTTTATATGTA 277  
 Oy 329 TGAAGCAGCTCGGGGGGTGAGTATTCGACTACCTGGTAAAGAGGGAGAGTACGC 388  
 Db 278 TGAATATGTCTGAGAGGTGAATATTTGACTACATCTGTAGCATGAGAGGGTTGAAG 337  
 Oy 389 CCAAGGAGCCCGAAAGTTCTCCGCCAGATGTGTCTGCGCTGGACTCTGCGACAGCT 448  
 Db 338 AGATGGAAGCCGAGCGGCTCTTTCAGAGATCTGTGCTGTGATTTACTGTCATAGGC 397  
 Oy 449 ACTCATCTGCGCAGAGAGACTTAAGCCGAGAACCTGCTTTTGGATGAGAAACAACA 508  
 Db 398 ATATGGTGTTCATCAGACCTGAAACAGAGAAATGTCTGTGGATGACACATATAAG 457  
 Oy 509 TCCGATTTGACAGCTTCGCGATGGCGTCCCTGACAGTGGGGAGACCCCTCTGGAGACA 568  
 Db 458 CCAAGATAGCCGATTTTGGATTTATCTAATATGATGTCAAGATGTTGAATTTCTGAGAACTA 517

Oy 569 GCTCGGGTCCCGCCCATATGCGTCTCCAGAGTGAATTAAGGGGAAAAATATGATGCC 628  
 Db 518 GTTCCGGATCTCAAAATTAATCAGACAGCTGAATCTCATCTGAGATGTGAGAGTC 577  
 Oy 629 GCCGGCAGACATGAGAGCTGTGAGATCATCCCTTCGCGCCCTGCTCGGGGGCTGCG 688  
 Db 578 CTGAAGTTGATATCTGAGCTGTGTGTATCTTATGCTCTTTTGTGGCACCTTCC 637  
 Oy 689 CTTTGTATGACAGACACTCCGCCAGCTGTGAGAGGTGAAGCGGCTCTTCACACA 748  
 Db 638 CATTTATGATGAGACATGTACTACGTATTTAAGAAGATCCGAGGGGTGCTTTTATA 697  
 Oy 749 TGCCCCACTTATCTCTCAAGTTCCAGAGCTCTCTGAGGGAATGATCGAATGGAGC 808  
 Db 698 TCCGAGATATCTCAATGTTCTGTGCGCACTCTCTGATGATATGCTGACAGGTTGACC 757  
 Oy 809 CCGAAGAAAGGCTCAGCTGGAGCAAAATTCAGAAACATCCTTGT 853  
 Db 758 CACTGAAGCGAGCAACTATCAAGACATTAAGAGACATGATGT 802

## RESULT 9

US-08-557-006C-36  
 ; Sequence 36, Application US/08557006C  
 ; Patent No. 6258547  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berl, Rajindar K.  
 ; APPLICANT: Carling, David  
 ; APPLICANT: Forster, Robert A.  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
 ; FILE REFERENCE: NGAP/PH37588/OST  
 ; CURRENT APPLICATION NUMBER: US/08/557,006C  
 ; CURRENT FILING DATE: 1996-03-06  
 ; PRIOR APPLICATION NUMBER: PCT/GB94/01093  
 ; PRIOR FILING DATE: 1994-05-20  
 ; PRIOR APPLICATION NUMBER: GB 9310489.1  
 ; PRIOR FILING DATE: 1993-05-21  
 ; PRIOR APPLICATION NUMBER: GB 9318010.7  
 ; PRIOR FILING DATE: 1993-08-31  
 ; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 36  
 ; LENGTH: 1783  
 ; TYPE: DNA  
 ; ORGANISM: Human AMP protein kinase  
 US-08-557-006C-36

Query Match 7.88; Score 182.6; DB 4; Length 1783;  
 Best Local Similarity 53.38; Pred. No. 7.8e-33;  
 Matches 408; Conservative 0; Mismatches 354; Indels 3; Gaps 1;

Oy 92 TGGCCCCCTATCGGCTGAGAGAGAGCTGGGCAAGAGACAGAGGGCTGTTAACTCG 151  
 Db 39 TCGGACACTAGCTGCTGGGCGACAGCTGGGCGTGGCGACCTTCGGCAAGAGAAATTG 98  
 Oy 152 GGGTCACCTGCATCAGGGTCAGAGAGTCCCATCAAGATCGTGAACCGGAGAAAGCTGT 211  
 Db 99 GAGAACATCAATTAACAGGCCATTAAGTGGCAGTTAAATCTTAATAGACAGAAATTC 158  
 Oy 212 CGAGT---CGGTGCTGATGAAGGTGAGCGGAGATGCCATCTCTGAAGCTCATCGAAC 268  
 Db 159 GCACTTATAGATGTTGTTGAAAAAATAAAGCAAGAAATTCMAAATCTTAAACTCTTCGTC 218  
 Oy 269 ACCCAGATGCTCTCAAGCTCCACGAGCTACAGAGAACAAATATTTGTAACCTGGTTC 328  
 Db 219 ATCTCATATATTAACAATAACAGAGTATGATGACATCCCAACAGATTTTATATGTA 278  
 Oy 329 TGAAGCAGCTCGGGGGGTGAGTATTCGACTACCTGGTAAAGAGGGAGAGACTGACGC 388  
 Db 279 TGAATATGTCTGAGAGGTGAATATTTGACTACATCTGTAAAGCATGAGACGGTTGAAG 338  
 Oy 389 CCAAGGAGCCCGAAAGTTCTCCGCCAGATGTGTCTGCGCTGGACTCTGCGACAGCT 448

```

Db 339 AGATGAGACCGGCGCTTTTACAGAGATTCTGTCTGTGATTAATCTCATAGCC 398
QY 449 ACTCCATCTGCGACAGAGACCTTAAAGCCGAGAACCTGTTTGGATGAGAAAAACACA 508
Db 399 ATATAGTTGTTTATCTGAGAGACTGAAGACAGAAATGCTCTGTGATGACATGATG 458
QY 509 TCCGATTCAGACATCTGGCAGTGGCGTCCCTGCAGGTGGGGGACACCTCTCGAGACCA 568
Db 459 CCAAGATAGCCGATTTGGATTATATATATGATGTCAGATGTGAATTTCTGAGAACTA 518
QY 569 GCTGCGGCGTCCCCCATTTATGCTGTCCAGAGGTGATTAAGGGGAAAAATATGATGGCC 628
Db 519 GTTCCGATCTCCAAATTAATGACAGACCTGAAGATATCTCAGGCAATTTATGAGGTC 578
QY 629 GCGGCGACAGATGTGAGCTGTGAGTCACTCTTCCGCTGCTGCGGGGCTCTGC 688
Db 579 CTGAAGTTGATATCTGAGAGCTGTGTTTATCTTGTATGCTTTTGTGTGACACCTCC 638
QY 689 CCTTGTATGACAGACCTCCGCACTGCTGAGAGAGTGAACGGGCGCTCTCCACA 748
Db 639 CATTGTATGATGACATGATCTACCTAGTTATTAAGAAATCCGAGGGGTGCTTTTATA 698
QY 749 TGGCCCACTCATCTCCAGATTCAGAGCCTCTGAGGGGAAATGATCGAAGTGAGC 808
Db 699 TCCAGATATCTAAATCTGCTGTCTGCCACTCTCTGATGATATGCTGCGAGTTGACC 758
QY 809 CCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCCTTGT 853
Db 759 CACTGAACGAGCAATATCAAGACATAGAGACATGAATGT 803

```

## RESULT 10

```

US-08-677-298-1
; Sequence 1, Application US/08677298
; Patent No. 5863729

```

## GENERAL INFORMATION:

```

APPLICANT: Pimica-Worms, Helen
TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN TCAK-1
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: USA
ZIP: 80303

```

## COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/677,298
FILING DATE: 09-JUL-1996
CLASSIFICATION: 536

```

## ATTORNEY/AGENT INFORMATION:

```

NAME: Caruthers, Jennie M.
REGISTRATION NUMBER: 34,464
REFERENCE/DOCKET NUMBER: 9-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089

```

## INFORMATION FOR SEQ ID NO: 1:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 2698 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: not relevant
MOLCULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: 376..2565
; US-08-677-298-1

```

```

Query Match
Best Local Similarity 7.6%; Score 177.8; DB 2; Length 2698;
Matches 404; Conservative 0; Mismatches 377; Indels 0; Gaps 0;

```

```

QY 92 TGGGCCCCCTTCGCGTGGAGAAAGCGTGGGCAAGAGACAGACAGCGCTGTTAACTCG 151
Db 533 TCGGAAGACTACAGACTTGTGAAAACATCGGCAAGGGAAATTTGCAAAATGAAATGG 592
QY 152 GGGTCCAGCTGATCAGCGGGTACAGAGTCCCATCAAGATCGTAACCGGGAAGAGCTGT 211
Db 593 CAAGACTATATCTTACAGGACAGAGAGTGGCAATAAAAATATATGACAAAATCAGTTGA 652
QY 212 CGGAGTGGGTGATGATGAGTGGAGGGGAGATCCGATCTGTAAAGTATCGAAGACC 271
Db 653 ATCCAAAGCTTACAAAAGCTCTTCAGAGAGATAGAAATATATGAAATTTAAATCATC 712
QY 272 CACATGCTCAAGCTCCAGAGCTTACAGAAACAGAAATATTTGACCTGGTCTGG 331
Db 713 CCAATATAGTGAAGTTATTTGAAATCATTTGAATGAAAAACACTTCACTATATATG 772
QY 332 AGCAGCTTCGGGGGTGAGCTATTTGACTACTGTAAGAAAGAGGAGACTGACGCCA 391
Db 773 AATATGCAAGTGGAGGAGATATTTGATTTGTTGTCACATGGCGAGATGAAGAAA 832
QY 392 AGGAGCGCGGAAGTTCTTCCGCAAGATTTGCTGCTGGCTGAGACTTCTGCCAGACTACT 451
Db 833 AAGAGCAAGATCTAAATTTAGACAGATGTGCTGCAATTCATCTCAGATCAGAAAC 892
QY 452 CCATCTGCCACAGAGACCTTAAAGCCGAGAACTGCTTTTGGATGAGAAAAACAATCC 511
Db 893 GGAATCGTACATGAGACCTGAAAGCTGAAATATTTGTTGATGCCGATTTGAACATTA 952
QY 512 GCATTCGAGACTTCGGCATGCGCTCCCTGCAAGTGGGGGAGCAGCCTCTGAGACCACT 571
Db 953 AAATAGCAGATTTGGTTTAAAGCAATGATTTACTGTTGGGGTAAACTGCACAGTTT 1012
QY 572 GGGGTCGCCCAATTATGCTGTCCAGAGTGAATTAAGGGGAAAAATATATGCGGCC 631
Db 1013 GTGGCAGTCTCCATAGACAGACCTGAGCTCTTCAGGGGCAAAATATACGGCCAG 1072
QY 632 GGGCAGCATGTGAGAGCTGAGAGTCACTCTTCGCCCTCTGATGGGGCTTGCCCT 691
Db 1073 AAGTGAATGTGTGAGAGCTGGGGGTCAATTTATACACTATGATGCTCACTTCCCT 1132
QY 692 TTGATGACGAAACCTTCGCCAGCTGCTGAGAGAGTGAAGCGGGCGCTTCCACATGC 751
Db 1133 TTGATGGGCAAAACCTTAAGAACTGAGAGAGATATTAAAGGGAATTAACGAATTC 1192
QY 752 CCCACTTCAATCTCCAGATTTGCCAGAGCTCTCGAGGGGAATATGAAATGAGCCG 811
Db 1193 CTTTACATGTCTACAGAGCTGTGAACCTTCGAAACGTTTCTGCTGCTAAATCCAA 1252
QY 812 AAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTCTGTACTAGCGGGAACAG 871
Db 1253 TTAAGCGGCACTCTAGAGCAAAATCATGAAGAGACAGTGTGATCAATGACGGGATGAG 1312
QY 872 A 872
Db 1313 A 1313

```

## RESULT 11

```

US-08-557-006C-44
; Sequence 44, Application US/08557006C
; Patent No. 6258547

```

## GENERAL INFORMATION:

```

APPLICANT: Beri, Rajindar K.
APPLICANT: Carling, David
APPLICANT: Forder, Robert A.

```

TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
 FILE REFERENCE: NCAP/PM37588/UST  
 CURRENT APPLICATION NUMBER: US/08/557,006C  
 CURRENT FILING DATE: 1996-03-06  
 PRIOR APPLICATION NUMBER: PCT/GB94/01093  
 PRIOR FILING DATE: 1994-05-20  
 PRIOR APPLICATION NUMBER: GB 9310489.1  
 PRIOR FILING DATE: 1993-05-21  
 PRIOR APPLICATION NUMBER: GB 9318010.7  
 PRIOR FILING DATE: 1993-08-31  
 NUMBER OF SEQ ID NOS: 44  
 SOFTWARE: Patentln Ver. 2.1  
 SEQ ID NO 44  
 LENGTH: 1747  
 TYPE: DNA  
 ORGANISM: Human  
 FEATURE:  
 NAME/KEY: gene  
 LOCATION: (1)..(1747)  
 OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -  
 OTHER INFORMATION: fragment begins at nucleotide 24 and ends with  
 US-08-557-006C-44

Query Match 7.5%; Score 175.2; DB 4; Length 1747;  
 Best Local Similarity 54.4%; Pred. No. 3.8e-31;

Matches 419; Conservative 0; Mismatches 343; Indels 8; Gaps 3;

92 TGGGCCCCCTAGCGCTGAGAGAGAGCGTGGGCAAGAGACAGAGCGGTGTTAACTCG 151  
 38 TCGGACACTACGCTGCTGGGAGACACCCCTGGCGTCCGACCTTCGGCAAGTGAAGATG 97  
 152 GGGTCCACTGACGAGGCTGAGAGGTCGCCATCAGATGCTGTAACCGGAGAACTGT 211  
 98 GAGAAACATCAATTCAGAGCGCATTAAGTGCAGTTAAGATCTTAATAGACAGAGATTC 157  
 212 CGGAGT---CGGTCTGATGAGAGGTGAGCGGAGATGCCATCCTGAACCTCAACGAA 268  
 158 GCAATTGATGTTGTTGAGAAATTAAGAGAAATTCATAATCTTTCGTC 217  
 269 ACCACATGTCCTCAGCTCCAGCAGCTCTACGAGAACAAATATTTGACTGTTTC 328  
 218 ATCTCATATTAATCAAACTCAACAGTGAAGACAGCTCAACAGACTTTTATGATGA 277  
 329 TGAACACGCTCTGGGGGCTGAGCTATTGACTACCTGTTAAGAAAGGAGAGACTGACG 388  
 278 TGGAAATATGTTCTGAGAGTGAATGTTGACTCATCTGTAACACGAGGAGGTGAAG 337  
 389 CCAAGGAGGCCGAAAGTTCTCCGCCAGATTTGTTGCGCTGAGCTTCG--CCACAG 446  
 338 AGGTGAAGCTCGCGGCTCTCCAGCAGATTCGTGCGGTGAGTACTGCTCAGAG 397  
 447 CTACTCCATCTGCCACAGAGACTTAAGCCCGAGAACCTGCTTTTGATGAGAAAAACA 506  
 398 GCACATGTTGTCACAGGAGCTGAAGCCAGAGAACTGTTGCTGAGGCCCAAGATGAA 457  
 507 CATCCGATTCAGACATTCGGCATGGCGTCCCTGCAGGTGGGGGACAGCCTCTGAGAG 566  
 458 TGTAAATATGACTGACTCGGAGCTCTTAATATGATGTCAGATGTTGATTTCTAGAC 517  
 567 CAGGTGGGGGCCCCCATTTATGGTGTCCAGAGGTGATTAAGGGGGGAAAAATATGATG 626  
 518 TAGTGTGAGTGGCCAAATTTGACAGCAGCGAGGTCAITTCAGAAAGGCTGATGGG 577  
 627 CCGCGGAGGACACATGTGAGCTGTGAGTATCTTCGCTTCGCTGCTGGGGGCTCT 686  
 578 TCCGAGAGTGTATCTGAGACTGTGTGTATCTCTGTATGCCCTTCTGTGAGCCCT 637  
 687 GCCCTTTGATGAC---GACAACTCCGCCAGCTGTGAGAAAGGTGAACGGGGCTCTT 743  
 638 CCGCTTGACAGACATGAGCAGCTGCTTAAAGATCCGAGGGGGTGTGTT 697  
 744 CCAATGCGCCACATTCATTCCTCCAGATTGCCAGAGCCTCTGAGGGGAAATGATGAAAT 803

Db 698 CTACATCCCGAGTATCTCAACCTTCTATTTGCCACTCTGCTGATGACATGCTGCAGCT 757  
 804 GGAGCCCGAAAAAGGCTCAGTGTGAGCAAAATTCAGAAACATCTCTGT 853  
 758 GAGCCCTTGAAGGACCACTATCAAAAGACATACGAGAGATGAATGCT 807

RESULT 12

US-09-930-181-5  
 Sequence 5, Application US/09930181  
 Patent No. 6455292

GENERAL INFORMATION:

APPLICANT: Origene Technologies

TITLE OF INVENTION: Full-Length Serine Protein Kinase In Brain and Pancreas

FILE REFERENCE: 16U 101 VI

CURRENT APPLICATION NUMBER: US/09/930,181

NUMBER OF SEQ ID NOS: 18

SOFTWARE: Patentln version 3.0

SEQ ID NO 5

LENGTH: 213

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(213)

US-09-930-181-5

Query Match 6.2%; Score 145; DB 4; Length 213;  
 Best Local Similarity 86.5%; Pred. No. 1.7e-24;

Matches 160; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

74 CCCAGCAGCCCATATATGTGGGCCCTATCGCTGAGAGACGCTGGCAAGAGACAGA 133  
 29 CGCAGACAGCGGAGTATGTTGGGCTTACCGCTGAGAAACGCTGGCAAGGCGCAGA 88  
 134 CAGGCTGTGTTAACTGGGGTCCACTGATCAGCGGTCAAGAGTCCCATCAAGATCG 193  
 89 CAGGTGTGTAAGCTGGGGTTCAGTGCCTGACCTGCCAAGAGTGGCATCAAGATCG 148  
 194 TGAACCGGAGAAAGCTTCGAGTCTGCTGATGAGAGGTGAGGCGGAGATCGCATCC 253  
 149 TCAACCTGTAAGAGCTAGCAGATCGGTGATGATGAAGTGGAGCGGAGATCGCATCC 208  
 254 TGAAG 258  
 209 TGAAG 213

RESULT 13

US-09-159-385-3  
 Sequence 3, Application US/09159385  
 Patent No. 5958748

GENERAL INFORMATION:

APPLICANT: AKIRA, SHIZUO

TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE

FILE REFERENCE: PH-569

CURRENT APPLICATION NUMBER: US/09/159,385

EARLIER APPLICATION NUMBER: JP97/261589

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 3

LENGTH: 2132

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (94)..(1455)

US-09-159-385-3

Query Match 5.88; Score 135.6; DB 2; Length 2132;  
 Best Local Similarity 54.1%; Pred. No. 4.7e-22;  
 Matches 360; Conservative 0; Mismatches 279; Indels 27; Gaps 3;

```

QY 78 GCACGCCCAATATGTGGGCCCCCTATCGGCTGGAAGAAGACGCTGGGCAAGAGCAACAGG 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 GCAGAGGAGAGCTGGAGACCATATATGAGATGGGGAGAGCTGGGACAGCGGCAATTC 167
QY 138 GCTGTAAACTCGGGGTCCACATCAGCGGCTCAGAAAGGTGGCCATCAAGATCGTGA 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 GATCTGCGGAAGTCCGGGAGAAAGGACAGGAGATACGACGACAGTTTATCAA 227
QY 198 CCGGAGAGACTGTC-----GAGTCGCTGTGATGAAGTGAAGTGAAGGAGAT 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 GAAGCGCGCTGTATCATCCAGCCGCGTGGGGTGAAGCGGAGAGATCGAGCGGAGGT 287
QY 246 CGCATCTCTGAAGCTATCGAACACCCACATGTCTCAAGCTCCAGAGCTTACGAGAA 305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 GAACATCTCTCGGAGATCCGGACCCCAACATCATCCTTCACGACATCTTCTGAGAA 347
QY 306 CAAGAAATATTTGATCTGTTCTGGAGCAGCTCGGGGGGTGAGCTATTGAGTACT 365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 CAAGAGGAGAGCTGTCTCTCATCTGAGAGCTGTCTCTGCGGGAGCTTTGACTTCT 407
QY 366 GGTAAAGAGGGAGAGACTGACGCCCAAGAGAGCGCCGAAAGTTCTTCCGACAGATTGTCTC 425
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 408 GCGGAGAGAAAGATCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 467
QY 426 TGGGCTGAGCTTGTGACAGAGCTACTCATCTGACAGAGAACCTTAAGCCCGAAGACT 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 468 CGCGCTTCACTACTGCTGATCTTAAGCGATCGACACTTTTACAGAAACCGGAAACAT 527
QY 486 GCTTTGGATGAGAAAAAC-----ACATCCGATTTGAGAGACTTGGAGTGGC 533
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 528 CATGCTGTGAGACAAAGACGTCGCCAACCCAGATCAAGCTCATGCACTTGGAGTGGC 587
QY 534 GTCCCTGACAGGTGGGAGACAGCTCTGAGAGACAGAGCTGCGGGTCCCCCATTTATGCTG 593
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 588 GCACAAGATCGAGCGGGGAGAACGATTCAAGAACATCTTCCGACACCCGAGTTGTGGC 647
QY 594 TCCAGAGGTATTAAGGGGAAAAATATGATGGCGCGGCGGACAGATGAGAGCTGTGG 653
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 648 CCCAGAGATTTGTAACTATGAGCCGCTG---GGCTTGAGAGCGGACATGTGAGAGCTGG 704
QY 654 AGTCATCTCTTCCCTGCTGTGGGGGCTGTGCTTGTGATGAGAGCAACCTCCGCCA 713
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 705 TGTATCATCACTATATCTCTCTGAGCGGTGATCCCGTTCTCTGGGAGAGACCAAGCAGGA 764
QY 714 GCTGCT 719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 765 GACGCT 770

```

RESULT 14  
 US-09-186-277-3

```

; Sequence 3, Application US/09186277
; Patent No. 6171841
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: 081356/0128
; CURRENT APPLICATION NUMBER: US/09/186, 277
; CURRENT FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2132
; TYPE: DNA
; ORGANISM: Homo sapiens

```

FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (94)..(1455)  
 US-09-186-277-3

Query Match 5.88; Score 135.6; DB 4; Length 2132;  
 Best Local Similarity 54.1%; Pred. No. 4.7e-22;  
 Matches 360; Conservative 0; Mismatches 279; Indels 27; Gaps 3;

```

QY 78 GCACGCCCAATATGTGGGCCCCCTATCGGCTGGAAGAAGACGCTGGGCAAGAGCAACAGG 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 GCAGAGGAGAGCTGGAGACCATATATGAGATGGGGAGAGCTGGGACAGCGGCAATTC 167
QY 138 GCTGTAAACTCGGGGTCCACATCAGCGGCTCAGAAAGGTGGCCATCAAGATCGTGA 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 GATCTGCGGAAGTCCGGGAGAAAGGACAGGAGATACGACGACAGTTTATCAA 227
QY 198 CCGGAGAGACTGTC-----GAGTCGCTGTGATGAAGTGAAGTGAAGGAGAGAT 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 GAAGCGCGCTGTATCATCCAGCCGCGTGGGGTGAAGCGGAGAGATCGAGCGGAGGT 287
QY 246 CGCATCTCTGAAGCTATCGAACACCCACATGTCTCAAGCTCCAGAGCTTACGAGAA 305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 GAACATCTCTCGGAGATCCGGACCCCAACATCATCCTTCACGACATCTTCTGAGAA 347
QY 306 CAAGAAATATTTGATCTGTTCTGGAGCAGCTCGGGGGGTGAGCTATTGAGTACT 365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 CAAGAGGAGAGCTGTCTCTCATCTGAGAGCTGTCTCTGCGGGAGCTTTGACTTCT 407
QY 366 GGTAAAGAGGGAGAGACTGACGCCCAAGAGAGCGCCGAAAGTTCTTCCGACAGATTGTCTC 425
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 408 GCGGAGAGAAAGATCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 467
QY 426 TGGGCTGAGCTTGTGACAGAGCTACTCATCTGACAGAGAACCTTAAGCCCGAAGACT 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 468 CGCGCTTCACTACTGCTGATCTTAAGCGATCGACACTTTGAGTGAAGCCGAAACAT 527
QY 486 GCTTTGGATGAGAAAAAC-----ACATCCGATTTGAGAGACTTGGAGTGGC 533
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 528 CATGCTGTGAGACAAAGACGTCGCCAACCCAGATCAAGCTCATGCACTTGGAGTGGC 587
QY 534 GTCCCTGACAGGTGGGAGACAGCTCTGAGAGACAGAGCTGCGGGTCCCCCATTTATGCTG 593
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 588 GCACAAGATCGAGCGGGGAGAACGATTCAAGAACATCTTCCGACACCCGAGTTGTGGC 647
QY 594 TCCAGAGGTATTAAGGGGAAAAATATGATGGCGCGGCGGACAGATGAGAGCTGTGG 653
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 648 CCCAGAGATTTGTAACTATGAGCCGCTG---GGCTTGAGAGCGGACATGTGAGAGCTGG 704
QY 654 AGTCATCTCTTCCCTGCTGTGGGGGCTGTGCTTGTGATGAGAGCAACCTCCGCCA 713
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 705 TGTATCATCACTATATCTCTCTGAGCGGTGATCCCGTTCTCTGGGAGAGACCAAGCAGGA 764
QY 714 GCTGCT 719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 765 GACGCT 770

```

RESULT 15  
 US-09-159-385-4

```

; Sequence 4, Application US/09159385
; Patent No. 5958748
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: PH-569
; CURRENT APPLICATION NUMBER: US/09/159, 385
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0

```

SEQ ID NO 4  
LENGTH: 1429  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (10)..(1353)  
US-09-159-385-4

Query Match 5.6%; Score 130.6; DB 2; Length 1429;  
Best Local Similarity 56.5%; Pred. No. 5.8e-21;  
Matches 291; Conservative 0; Mismatches 209; Indels 15; Gaps 2;

QY 223 CTGATGAAGGTGAGCGGGAGTCGCCATCTGAGCTCATGAAACACCAATGCTCTC 282  
DB 181 CGGAGAGAGATCGAAGCGCGAGTCGATCTCGGCGAGATCCGCCAACATCATATA 240  
QY 283 AAGCTCCAGAGCTGACGAAACAAATTTTACCTGCTTCTGAGACAGCTCG 342  
DB 241 ACACCTGATGACGTGTGGAACAAAGACAGATGTGTGATCTGAGCTGTGTCC 300  
QY 343 GGGGTGAGCTTATGACCTGCTGTAAGAAGGGAGACTGACGCCCAAGAGGCCCGA 402  
DB 301 GGTGGGAGCTTTTCACCTCTGCGCGAGAGAGTCAATTGACGAGATGAGGCCAG 360  
QY 403 AAGTCTTCGCGAGATGTGTGCTGCGCTGACTTTCGCCACAGCTACTCCATCTGCCAC 462  
DB 361 CAGTCTCAAAACAAATCTAGACGGGTGCTCACTGACCTGCACTCCAAAGCGCATGCGACAC 420  
QY 463 AAGAGCTTAAAGCCGAAACCTGCTTTTGGATGAGAAAAACAATCCGC----- 513  
DB 421 TTTGACCTGAAGCCGAAACATCATGTGTGGAACAAGCAGCAGCCAGCCCGCATT 480  
QY 514 ---ATTGACACTTGGCATGGCTCCCTGACGAGTGGGGACAGCCTCTGGAGACAGC 570  
DB 481 AAGCTATGACCTTTGGCATGGCGACAGATGAGGCTGGCAGCGAGTCAAGAACATC 540  
QY 571 TGGGGTCCCCCATTAATGCGGTGTCCAGAGGTATTAAGGGGAAAAATATGATGGCGC 630  
DB 541 TTTGGCACAACCCGAGTTTGTGCGCCCCCGAGATGTAATGAGCCA---CTTGGCTTG 597  
QY 631 CGGGCAGACATGAGAGCTGTGAGATCATCTTCCGCCCTGCTGTGGGGCTCTGCC 690  
DB 598 GAGGCTGACATGTGAGCATTTGGGCTCATCACTTCTGTGAGCGAGCGTCCCA 657  
QY 691 TTTGATGACGACAACCTCCGCGACAGCTGCGAGAA 725  
DB 658 TTCCTGGGCGAGACCAAGCAGAGACGCTGACGAA 692

Search completed: April 22, 2003, 00:26:08  
Job time : 66.0237 secs



GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 20:07:11 ; Search time 2126.25 Seconds  
(without alignments)  
17777.901 Million cell updates/sec

Title: US-10-003-690-3  
Perfect score: 2334  
Sequence: 1 atgtgtccggggcgaaga.....ccaaagggaccctctgcc 2334

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 33308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlun:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_estrc:\*  
9: gD\_estli:\*  
10: gD\_estl2:\*  
11: gD\_estl3:\*  
12: gD\_estl4:\*  
13: gD\_estl5:\*  
14: gD\_estl6:\*  
15: em\_estlun:\*  
16: em\_estlun:\*  
17: gD\_gss:\*  
18: gD\_gss:\*  
19: em\_gss\_hum:\*  
20: em\_gss\_hum:\*  
21: em\_gss\_vtl:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	759.8	32.6	889	14	B0434571 AGENCOURT
2	723.6	31.0	881	14	B0717903 AGENCOURT
3	651.2	27.9	831	12	BF529743 AGENCOURT
4	602.2	25.8	658	12	BF529743 AGENCOURT
5	592.4	25.4	698	14	BM948524 UI-M-EGOP
6	574.8	24.6	742	14	BM944329 UI-M-EGOP

7	543	23.3	671	14	B0178547
8	540.4	22.6	658	14	B0573641
9	527.6	22.2	836	9	AL538014
10	522.4	22.4	936	12	BF529630
11	510	21.9	510	14	B0086295
12	490.8	21.0	898	12	BF529174
13	490.2	21.0	600	14	B044032
14	487.8	20.9	753	12	BF796285
15	476.2	20.4	614	14	BM944236
16	460.4	19.7	728	14	BM944236
17	444	19.0	492	12	BF686432
18	440.4	18.9	597	13	BF686432
19	439	18.8	453	12	BF686432
20	429.4	18.3	733	14	B0178695
21	426.6	18.4	508	14	BM950539
22	415.4	17.8	521	14	B0442731
23	406.4	17.4	693	14	B0443644
24	406	17.4	961	14	W41136
25	404.4	17.3	551	12	B0715357
26	374.4	16.0	488	14	W60403
27	366.2	15.7	793	9	AL135353
28	365	15.6	596	10	BE222941
29	352.2	15.1	427	14	W18584
30	341.8	14.6	418	10	AM122891
31	338	14.5	712	14	B0180079
32	336.2	14.4	925	14	B0435328
33	331.8	14.2	609	14	BM949588
34	331.2	14.0	344	12	BF931102
35	326	14.0	367	12	BE763502
36	325	13.9	337	13	BM129669
37	319.8	13.7	394	14	W75642
38	318.2	13.6	1177	13	BM466629
39	307.2	13.2	733	12	BF395625
40	307.2	13.2	1882	11	BC017182
41	306	13.1	546	14	BM944658
42	305.2	13.1	563	10	BE251924
43	302.6	13.0	334	12	BF371253
44	293.2	12.6	740	12	BF222739
45	288.4	12.4	293	9	AA326737

## ALIGNMENTS

RESULT 1  
B0434571  
LOCUS  
DEFINITION  
AGENCOURT\_7917755 NIH\_MGC\_67 Homo sapiens CDNA clone IMAGE:6154749  
5', mRNA sequence.  
ACCESSION  
B0434571  
VERSION  
B0434571.1 GI:21173647  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens

REFERENCE  
1 (bases 1 to 889)  
AUTHORS  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished (1999)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: L1M13496 row: e column: 22  
High quality sequence stop: 681.  
Location/Qualifiers  
1..889

FEATURES  
source

BASE COUNT	207 a	246 c	274 g	160 t	2 others
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:6154749" /clone_1ib="NH.MGC.67" /issue_type="retinoblastoma" /lab_host="DH10B (phage-resistant)" /note="Organ: eye; Vector: PCMV-SPOrt6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."				

Query Match	32.6%	Score 759.8	DB 14	length 889
Best Local Similarity	98.5%	Pred. No. 2.7e-154		
Matches 788	Conservative	0	Mismatches 8	Indels 4
				Gaps 2

OY	104	GGCTGGAGAAAGACCGCTGGGGCAAAAGGACACAGACGGGCTGGTTAACTCGGGGTCACATCGCA	163
Db	9	GGGTGGAGAAAGACCGCTGGGGCAAAAGGACACAGGCGTGGTTAACTCGGGGTCACATCGCA	68
OY	164	TCACGGGTCAGAAAGGTCGCCATCAAGATCGTGAACCGGGAGAAAGCTGTGCGAGTCGGTGC	223
Db	69	TCACGGGTCAGAAAGGTCGCCATCAAGATCGTGAACCGGGAGAAAGCTGTGCGAGTCGGTGC	128
OY	224	TGATCAAGCTGGAGCGGGAGATCGGCATCTGTAACCTCATGGAACACCCACATGTCTCA	283
Db	129	TGATCAAGCTGGAGCGGGAGATCGGCATCTGTAACCTCATGGAACACCCACATGTCTCA	188
OY	284	AGCTTCACGAGCTGTACAGAAACAGAAATATTTGTACCTGGTTCTGGAGCACGTCCTGG	343
Db	189	AGCTTCACGAGCTGTACAGAAACAGAAATATTTGTACCTGGTTCTGGAGCACGTCCTGG	248
OY	344	GGGGTGACCTATTGCACTACTCGTTAAAGAAAGGAGACTGCAGCCCAAGAGGCCGAA	403
Db	249	GGGGTGACCTATTGCACTACTCGTTAAAGAAAGGAGACTGCAGCCCAAGAGGCCGAA	308
OY	404	AGTTCTTCGGCAGATTTGTGCTGGCGGAGACTTGTGCACAGTACATCTGCTCCACA	463
Db	309	AGTTCTTCGGCAGATTTGTGCTGGCGTGAATTTGTGCACAGTACATCTGCTCCACA	368
OY	464	GAGACCTAAAGCCCAAGAACCTGCTTTTGGATGAGAAAAACAACATCCGATTTGAGACT	523
Db	369	GAGACCTAAAGCCCAAGAACCTGCTTTTGGATGAGAAAAACAACATCCGATTTGAGACT	428
OY	524	TCGGCATGCGCTCCCTGCAAGGTGGGGGACAGCTCTCTGGAACACAGCTGCGGGTCCCCC	583
Db	429	TCGGCATGCGCTCCCTGCAAGGTGGGGGACAGCTCTCTGGAACACAGCTGCGGGTCCCCC	488
OY	584	ATTATGCGTGTCCACAGGTGATTAAAGGGGAAAAATATGATGGCCGCGGGCAGACATGT	643
Db	489	ATTATGCGTGTCCACAGGTGATTAAAGGGGAAAAATATGATGGCCGCGGGCAGACATGT	548
OY	644	GGAGCTGTGAGTCATCTCTTCCGCCCTGCTGCTGGGGGCTCTGCCCCCTTGTGATGACACA	703
Db	549	GGAGCTGTGAGTCATCTCTTCCGCCCTGCTGCTGGGGGCTCTGCCCCCTTGTGATGACACA	608
OY	704	AACCTCCGACAGTGTGTGGAGAAAGTGAACCGGGGGCTTCCACATGTGCCCATCTCATTC	763
Db	609	AACCTCCGACAGTGTGTGGAGAAAGTGAACCGGGGGCTTCTCCACATGTGCCCATCTCATTC	668
OY	764	CTCCAGATTGCGACAGCCCTCTGAGGGGAAATGATCGAAGTGGAGCCCAAAAAAGGCTCA	823
Db	669	CTCCAGATTGCGACAGCCCTCTGAGGGGAAATGATCGAAGTGGAGCCCAAAAAAGGCTCA	728
OY	824	GTCGTGAGCAAAATTCAGAAACATCTCTGG--TACCTTAGCGGGGGAACAACGAGCCAGACC	880
Db	729	GTCGTGAGCAAAATTCAGAAACATCTCTTGTATCACTAGAGGGGGGAACAACGAGCCAGACC	788
OY	881	CGTG--CTGAGACCAAGCCCC	899
Db	789	CGTGCCCTGAGACCAAGCCCC	808

[illegible]

Tissue Procurement: Dr. James R. Lupski  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIMD)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MSC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLMD at:  
<http://image.llnl.gov>  
Plate: LLM13580 row: c column: 11  
High quality sequence stop:650.

FEATURES	location/qualifiers
source	1. .881

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6186946"
/clone_id="Lupsk1_sympathetic_trunk"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/notes="vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-GCGACCCACCGCGCGCG-3' and
5'-GACGCTGCTTAGAGCGAGCGCGGCCCTT(15)-3', size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

```

	Query Match	Similarity	31.0%	Score	723.6	DB	14	Length	881
	Best Local	Similarity	95.3%	Pred. No.	1,9e+146				
	Matches	755	Conservative	0	Mismatches	36	Indels	1	Gaps
QY	485	TGCTTTTGGATGAGAAAAACAACATCCGATTCGACACTTCGGCATTCGGCGTCCCTGCAGG	544						
Db	1	TGCTTTTGGATGAGAAAAACAACATCCGATTCGACACTTCGGCATTCGGCGTCCCTGCAGG	60						
QY	545	TGGGGGACACGCTCCTCGGAGACACAGCTGGGGTCCCCCATTTATCGTGTCCAGAGGTGA	604						
Db	61	TGGGGGACACGCTCCTCGGAGACACAGCTGGGGTCCCCCATTTATCGTGTCCAGAGGTGA	120						
QY	605	TTTAAGGGGAAAAAATATGATTCGCGCGCGGGGACAGCATGTGAGATGTCATCTCT	664						
Db	121	TTTAAGGGGAAAAAATATGATTCGCGCGCGGGGACAGCATGTGAGATGTCATCTCT	180						
QY	665	TGCGCCCTGCTCGTGGGGGCTTCGCCCTTTGATGAGACAGACAACCTCCGCCAGCTGCTGGAGA	724						
Db	181	TGCGCCCTGCTCGTGGGGGCTTCGCCCTTTGATGAGACAGACAACCTCCGCCAGCTGCTGGAGA	240						
QY	725	AGGTGAACCGGGGCTCTTCCATATGCCCACTTCATTCCTCCAGATTGCCAGAGCCCTTC	784						



KEYWORDS EST.  
SOURCE Homo sapiens  
ORGANISM Human  
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
TITLE NIH-MGC http://mgi.nci.nih.gov/  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Inceye Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL0858 row: n column: 09  
High quality sequence stop: 612.  
Location/Qualifiers  
1. 658  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4932680"  
/clone\_lib="NCI-CGAP\_Brn67"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q  
loss"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1; NotI;  
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2.3 kb. Constructed by Life  
Technologies. Note: this is a NCI-CGAP library."  
BASE COUNT 145 a 223 g 96 t  
ORIGIN  
Query Match 25.4%; Score 602.2; DB 12; Length 658;  
Best Local Similarity 98.4%; Pred. No. 3.7e-120;  
Matches 619; Conservative 0; Mismatches 8; Indels 2; Gaps 1;  
QY 859 GGGGGGAAACAGCAGCCGCTGCTGAGCAGCCCTGCGCGCGGAGTGCCTG 918  
DB 1 GGGGGGAAACAGCAGCCGCTGCTGAGCAGCCCTGCGCGCGGAGTGCCTG 60  
QY 919 CGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 978  
DB 61 CGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
QY 979 GGGTCTGAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1038  
DB 121 GGGTCTGAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
QY 1039 GAAAGAT 1098  
DB 181 GAAAGAT 240  
QY 1099 CAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1158  
DB 241 CAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
QY 1159 CAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1218  
DB 301 CAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 1219 GATGCGGG 1278  
DB 361 GATGCGGG 420  
QY 1279 AGCCAGAGATCCCGTAGCGTCACTGAGAGCTTCAAGGCTGCTTCAAGC 1338  
DB 421 AGCCAGAGATCCCGTAGCGTCACTGAGAGCTTCAAGGCTGCTTCAAGC 480  
QY 1339 AGCCAGAGATCCCGTAGCGTCTTTTCTTTTCAAGGAGCGGGGGGGTGA 1398

DB 481 AGCCAGAGATCCCGTAGCGTCTTTTCTTTTCAAGGAGCGGGGGGGTGA 540  
QY 1399 GGG 1458  
DB 541 GGG 598  
QY 1459 GGG 1487  
DB 599 GGG 627  
RESULT 5  
LOCUS BM948524  
DEFINITION UI-M-EG0P-bve-1-10-0-UI.r1 NIH\_BMAP\_EG0P Mus musculus CDNA clone  
IMAGE:5691033 5', mRNA sequence.  
ACCESSION BM948524  
VERSION BM948524.1 GI:19432115  
KEYWORDS EST.  
ORGANISM Mus musculus  
SOURCE house mouse.  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
TITLE NIH-MGC http://mgi.nci.nih.gov/  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Dr. James Lin, University of Iowa  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)  
Seq primer: pyx-5.  
Location/Qualifiers  
1. 698  
/organism="Mus musculus"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5691033"  
/clone\_lib="NIH\_BMAP\_EG0P"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 18.5 dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/note="Organ: brain; Vector: pyx-Asc; Site: 1; EcoR I;  
Site: 2; Not I; The library was constructed according to  
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured mRNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with an  
oligo-dT primer containing a Not I site. Double stranded  
cDNA was size selected according to mRNA size fraction,  
ligated with EcoR I adaptor, digested with Not I, and then  
cloned directionally into pyx-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
, is CAGCCAGAC. This library was created for the  
University of Iowa Mouse Brain Molecular Anatomy Project  
(BMAP). 'Gene Discovery in the Developing Mouse Nervous  
System', supported by National Institute of Mental Health  
(NIMH), Hemlin Chin, Ph.D., program coordinator."  
BASE COUNT 165 a 190 c 209 g 134 t  
ORIGIN  
Query Match 25.4%; Score 592.4; DB 14; Length 698;  
Best Local Similarity 90.5%; Pred. No. 5e-118;  
Matches 632; Conservative 0; Mismatches 66; Indels 0; Gaps 0;  
QY 332 AGCAGCTCTGGGGGGGGTGAAGTATTCAGTACCTGTAAAGAGGAGAGTACGCCCA 391



Query Match	Best Local Similarity	Score	DB	Length
23.3%	90.0%	543	14	671
		Pred. No. 2.6e-107		

  

FEATURES	SOURCE
Seq primer: PYX-5.1	Location/Qualifiers
1. 671	/organism="Mus musculus"
	/strain="C57BL/6"
	/db_xref="taxon:10090"
	/clone="IMAGE:5701154"
	/clone_lib="NIH_BMAP_EVO"
	/tissue_type="whole brain"
	/dev_stage="embryo 15.5 dpc"
	/lab_host="DH10B (T1 phage resistant)"
	/note="Organ: brain; Vector: pYX-Asc; Site: 1; Ecor I; site: 2; Not I. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTCGCTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemlin Chih, Ph.D., program coordinator."
BASE COUNT	160 a 188 c 200 g 122 t 1 others
ORIGIN	

  

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 671)	NIH-MGC <a href="http://mgi.nci.nih.gov/">http://mgi.nci.nih.gov/</a> .	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
	Contact: Robert Struhsberg, Ph.D.	Email: <a href="mailto:cgaeb@mail.nih.gov">cgaeb@mail.nih.gov</a>		
	Tissue Procurement: Dr. James Lin, University of Iowa			
	cDNA Library preparation: Dr. M. Bento Soares, University of Iowa			
	cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa			
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa			
	cDNA Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLM at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>			
	This clone was contributed by the Brain Molecular Anatomy Project (BMAP)			

  

RESULT 7	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM
B0178547	671 bp	mRNA	linear	EST 30-APR-2002		
UI-M-EVO	-b-ws-b-03-0-ui.r1	NIH_BMAP_EVO	Mus musculus	cDNA clone		
IMAGE:5701154	5'	mRNA sequence.				
B0178547	1	GI:20354039				
EST.						
house mouse.						
Mus musculus						
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						

  

Db	542	GCTGTTGAGACACCGCGGCTGTCACAGAGAACTGCGAAGCGAGGAGAAACCAAG	601
0y	1040	AAAAGATGATATATATATCTGC-TTTTGGATCGGAAGAGCGGATATCCAGCTGTAGAGC	1098
Db	602	AAAAGATGATATATATATATTTGCTTTTGGATCGGGAAGACGGGTATCTGATGAAGC	661
0y	1099	CAGGACCTGACCCCTCCCGGAAATGATGTTGACC-CCCCCGGAAGCGTGTGATTCGCCAT	1157
Db	662	CAGGA-CTGCCCTCTCGGAATGATGTGACCACTCCGTAGCGGTGTGATTTCCCCCAT	720
0y	1158	GCTGAGCCGTCACGGGAACGG	1179
Db	721	GCTGAGCANACACGGAACCG	742

[illegible]

Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:

http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: PYX-5.

# FEATURES

## source

Location/Qualifiers  
1. 658  
/organism="Mus musculus"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5718158"  
/clone\_lib="NIF BMAP.FDO"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 12.5 dpc"  
/lab\_host="DHI0B (T1 phage resistant)"  
/note="Organ: brain; Vector: pyx-asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

## BASE COUNT

153 a 185 c 195 g 125 t

## ORIGIN

Query Match 23.2%; Score 540.4; DB 14; Length 658;  
Best Local Similarity 89.8%; Pred. No. 9.4e-107;  
Matches 591; Conservative 0; Mismatches 66; Indels 1; Gaps 1;

QY 735 GGGGCTCTCCACATGCCCCCATTCCTCCAGATTCGACAGAGCTCTCTGAGGGGAAT 794  
DB 2 GGGGGCTCTCCACATGCCCCCATTCCTCCAGATTCGACAGAGCTCTCTGAGGGGAAT 61  
QY 795 GATGAGGTGAGGCCGGAAGAGGCTGAGTGGAGCAATTCGAAACATCTCTGGTA 854  
DB 62 GATTGAGGTGAGGCCGGAAGAGG-TCAGTCTGAGCAAAATTCGAAACATCTCTGGTA 120  
QY 855 CCTAGAGCGGGAACAGCAGACGACCCGCTGAGAGCCAGCCCGCGCGGTAGC 914  
DB 121 TCTGGGGGGGGAACAGCAGACGACCCGCTGAGAGCCAGCCCGCGCGGTAGC 180  
QY 915 CATGGGAGCTGCCATTCACAGAGAGCTGGAGCCCGACGCTCTAGAGAGCATGCGATC 974  
DB 181 CATGCTTGTGCTGCTTCCATGAGCGAGCTGGAGCCCGACGCTCTAGAGAGCATGCGATC 240  
QY 975 ACTGGGCTGCTTCCAGAGAGCGGAGAGGCTGATCGGAGCTGCGAGAGAGAGAGAA 1034  
DB 241 TCTGGGCTGCTTCCAGAGAGCGGAGAGGCTGATCGGAGCTGCGAGAGAGAGAGAA 300  
QY 1035 CCAAGAAAGATGATATATATATCTCTTTGATGAGAGAGAGAGAGAGAGAGAGAG 1094  
DB 301 CCAAGAAAGATGATATATATATCTTTTGTGATGAGAGAGAGAGAGAGAGAGAGAG 360  
QY 1095 GGACGAGAGCTGCTCCCGGAGATGATGTTGACCCCGCGGAGAGCGTGTGAGATTCTCC 1154  
DB 361 AGACGAGAGCTGCTCCCGGAGATGATGTTGACCCCGCGGAGAGCGTGTGAGATTCTCC 420  
QY 1155 CATGCTGAGCGCTGAGAGAGCGGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1214  
DB 421 CATGCTGAGCGAG 480  
QY 1215 CACGAG 1274  
DB 481 CACGAG 540

QY 1275 GCACAGCCAGAGATCCCTAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1334  
DB 541 GCACAGCTGAGAGATCCCGACGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
QY 1335 AAGCAGCCAG 1392  
DB 601 GAGCAGCCAG 658

## RESULT 9

## AL538014

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 836)  
I.L.W.B., Gruber, C., Jessen, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

## FEATURES

## source

Location/Qualifiers  
1. 836  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DF029YC21"  
/clone\_lib="LTL FL013.Fbrn1"  
/dev\_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"  
/lab\_host="DHI0B"  
/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@life.com URL : http://fulllength.invitrogen.com"

## BASE COUNT

181 a 230 c 268 g 152 t 5 others

Query Match 22.6%; Score 527.6; DB 9; Length 836;  
Best Local Similarity 81.4%; Pred. No. 5.9e-104;  
Matches 630; Conservative 4; Mismatches 136; Indels 2; Gaps 2;

QY 79 CACGCCAATATGTGGGCCCTATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 138  
DB 65 CACGCCAATATGTGGGCCCTATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 124  
QY 139 CTGTTAACTCGGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 198  
DB 125 CTGTTAACTCGGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 184  
QY 199 CGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 258  
DB 185 CGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 244  
QY 259 CTCATCGAAG 318  
DB 245 CTCATCGAAG 304  
QY 319 TACCTGTTGAG 378

```

|||||
Db 305 TACTGTGCTAGACACACGTGTGAGCTTTCGACTTACCTGGTGAAGAGGG 364
OY 379 AGACTGACGCCCAAGAGAGGCCGAAAGTTCCTCCGACATTTGTCTGCGCTGACATTC 438
Db 365 AGGCTGACGCCCAAGAGAGGCCGAAAGTTCCTCCGACATTTGTCTGCGCTGACATTC 424
OY 439 TGGCAGACGTCATCTGTCAGACACCTTAAAGCCCGAGAACCTGCTTTGTGATG 498
Db 425 TGGCAGACGTCATCTGTCAGACACCTTAAAGCCCGAGAACCTGCTTTGTGATG 484
OY 499 AAAAACAACATCCGATTCGACACTTCCGATGCGCTCCCTGACAGTGGGAGACAGCTC 558
Db 485 AAGAACAACATCCGATTCGACACTTCCGATGCGCTCCCTGACAGTGGGAGACAGCTC 544
OY 559 CTGGAGACCAAGCTGCGGCTCCCGCATTTATGCTGTCCAGAGGTATTAAGGGGAAAA 618
Db 545 TTGGAGACCAAGCTGCGGCTCCCGCATTTATGCTGTCCAGAGGTATTAAGGGGAAAA 604
OY 619 TATGATGCGCGCGCGAGACATGTCGAGCTGTGAGTCAATCTTGGCTGCTCTG 678
Db 605 TATGACGCGCGG-GAAGSGAGCTGTGAGCTGTGAGTCAATCTTGGCTGCTCTG 663
OY 679 GGGGCTGCTGCTTGTGATGACGACACCTCCGACGCTGTGAGAGGTGAAGAGGGGC 738
Db 664 GGGGCTGCTGCTTGTGATGACGACACCTCCGACGCTGTGAGAGGTGAAGAGGGGC 722
OY 739 GTCTTCCACATGCCCCACTTCTTCTCCAGATTCGACAGCTCTTGAAGGGAAATGATC 798
Db 723 GTCTTCCACATGCCCCACTTCTTCTCCAGATTCGACAGCTCTTGAAGGGAAATGATC 782
OY 799 GAAGTGAGCGCGGAAAAAGGCTGAGCTGTGAGCAATTCAGAAACATCTTGG 852
Db 783 GAGTGAGCGCGGAAAAAGGCTGAGCTGTGAGCAATTCAGAAACATCTTGG 836

RESULT 10
BF529630 936 bp mRNA linear EST 11-DEC-2000
LOCUS 602043855F1 NCL_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4181581
DEFINITION 5', mRNA sequence.
ACCESSION BF529630
VERSION BF529630.1 GI:11616993
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 936)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.lnl.gov
Plate: LAM9494 row: n column: 14
High quality sequence start: 8
High quality sequence stop: 568.
Location/Qualifiers
1. 936
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4181581"
/clone_id="NCL_CGAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"

```

```

/note="Organ: brain; Vector: pCMV-Sport6; Site: 1; NCI;
Site: 2; Salt; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCL_CGAP Library."
BASE COUNT 286 a 238 c 318 g 94 t
ORIGIN
Query Match 22.4%; Score 522.4; DB 12; Length 936;
Best Local Similarity 97.6%; Pred. No. 8.2e-103;
Matches 562; Conservative 0; Mismatches 11; Indels 3; Gaps 3;
OY 901 GGGCGCGGTAGCGATCGAGCGAGCTGCAATCCAGAGAGAGCTGGACCCGAGCTCTTA 960
Db 5 GGGCGCGGTAGCGATCGAGCGAGCTGCAATCCAGAGAGAGCTGGACCCGAGCTCTTA 64
OY 961 GAGAGCATGCGATCACTGGCTGCTTCAAGGACCGCGAGAGCTGCATCGGAGCTGGC 1020
Db 65 GAGAGCATGCGATCACTGGCTGCTTCAAGGACCGCGAGAGCTGCATCGGAGCTGGC 124
OY 1021 AGTGAGAGAGAGAGAGAGAGAGAGATATATATATCTCTTTTGTGATCGAGAGCGG 1080
Db 125 AGTGAGAGAGAGAGAGAGAGAGATATATATATCTCTCTTTTGTGATCGAGAGCGG 183
OY 1081 TATCCAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 184 TATCCAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 243
OY 1141 CGTGTGATTTCTCCATGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 244 CGTGTGATTTCTCCATGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 302
OY 1201 GAAGTCTGAGCATCAACCATGCTGCGGGGGGTGTGTGTCTTCTTGAACGACCCGAGG 1260
Db 303 GAAGTCTGAGCATCAACCATGCTGCGGGGGGTGTGTGTCTTCTTGAACGACCCGAGG 362
OY 1261 TTGAGATGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 363 TTGAGATGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421
OY 1321 TCTTCCAGCTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 422 TCTTCCAGCTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
OY 1381 GCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 482 GCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541
OY 1441 CGGGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1476
Db 542 TCGGGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 577

RESULT 11
B0086295 510 bp mRNA linear EST 29-APR-2002
LOCUS 1120b01.y1 Melton Normalized Human Islet 4 M-HIS 1 Homo sapiens
DEFINITION cDNA clone IMAGE:6135408 5' similar to TR:060843 060843 PUTATIVE
SERINE/THROMBIN PROTEIN KINASE ;, mRNA sequence.
ACCESSION B0086295
VERSION B0086295.1 GI:20045499
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 510)
Melton, D., Brown, J., Kenty, G., Permutt, A., Iac, C., Kaestner, K.,
Lemishka, I., Secombe, M., Brestelli, J., Gridwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Rittler, E., Ronko, I., Bennett, D., Cardenas
, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,
Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium

```

JOURNAL Unpublished (2000)  
 COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@molb.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:  
 Washington University Genome Sequencing Center this clone is  
 available royalty-free through LNLN; please contact the IMAGE  
 Consortium (info@image.llnl.gov) for further information  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 445.

FEATURES  
 source

1. 510  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_1lb="IMAGE:6135408"  
 /sex="Both"  
 /tissue\_type="Islets of Langerhans"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B"  
 /note="Organ: Pancreas; Vector: pSPORT1; Site\_1: Not 1;  
 Site\_2: Sal 1; Starting library constructed using  
 Superscript Plasmid Library Kit (Life Technologies). cDNA  
 made by oligo-dT priming. Size-selected by column  
 fractionation; average insert size 1.08 kb. Library was  
 amplified once on solid support and plasmid DNA from  
 library was prepared. The library DNA was normalized by  
 method #4 from Ronald, Lennon, and Soares 1996 Genome  
 Research 6:791-806; 0.5 microgram single-stranded library  
 plasmid DNA was mixed with 5 micrograms PCR product  
 representing library inserts and hybridized to an EcoT of  
 20. Single-stranded (unhybridized) plasmids were isolated  
 by hydroxyapatite chromatography and used to make this  
 library."

BASE COUNT 110 a 171 c 136 g 93 t

Query Match 21.9%; Score 510; DB 14; Length 510;

Best Local Similarity 100.0%; Pred. No. 3.4e-100; Mismatches 0; Indels 0; Gaps 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1692 TCGCTTACCGGCGGAGATGAGTCCCTACCGTGAAGAGATGTCACCTGACGCC 1751  
 |||||||  
 DB 1 TCGCTTACCGGCGGAGATGAGTCCCTACCGTGAAGAGATGTCACCTGACGCC 60  
 |||||||

OY 1752 AGAGTCTCCCGGAGTGCAGAAAGCTCTGTTGGGAACTTCATCTCTTGACAA 1811  
 |||||||  
 DB 61 AGAGTCTCCCGGAGTGCAGAAAGCTCTGTTGGGAACTTCATCTCTTGACAA 120  
 |||||||

OY 1812 AGAAGACAATATTCCTCTGCTAAAGACAACCTCTGACGATCAAGACGACAT 1871  
 |||||||  
 DB 121 AGAAGACAATATTCCTCTGCTAAAGACAACCTCTGACGATCAAGACGACAT 180  
 |||||||

OY 1872 CGTCATGCTTCTGTGATCCCAAGCTGATGACAGTGTGCTGACAGACCACTT 1931  
 |||||||  
 DB 181 CGTCATGCTTCTGTGATCCCAAGCTGATGACAGTGTGCTGACAGACCACTT 240  
 |||||||

OY 1932 CAGGGCGAGTAAAGGCGGAGGCGGCGCTCGTCTCCAAAAAGCCGCTTCCCA 1991  
 |||||||  
 DB 241 CAGGGCGAGTAAAGGCGGAGGCGGCGCTCGTCTCCAAAAAGCCGCTTCCCA 300  
 |||||||

OY 1992 GGTGACATCAGCTCTGAGGCTCAGAGCCCTCCCGGAGGAGGAGCGAGGAGG 2051  
 |||||||  
 DB 301 GGTGACATCAGCTCTGAGGCTCAGAGCCCTCCCGGAGGAGGAGCGAGGAGG 360  
 |||||||

OY 2052 TGTGTCATCTACTCGTCACTCTCANTCTGAGGCTCCAGCCGCTGCTTCAACG 2111  
 |||||||  
 DB 361 TGTGTCATCTACTCGTCACTCTCANTCTGAGGCTCCAGCCGCTGCTTCAACG 420  
 |||||||

OY 2112 AGTGGTGAACCATCCAGGACAGCTCTGAGCAGTCAATGACCAAGCCCTCCGACGC 2171  
 |||||||  
 DB 421 AGTGGTGAACCATCCAGGACAGCTCTGAGCAGTCAATGACCAAGCCCTCCGACGC 480  
 |||||||

OY 2172 CCTGGCAGAGAGAGAGAGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 2201  
 |||||||  
 DB 481 CCTGGCAGAGAGAGAGAGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 510  
 |||||||

RESULT 12 898 bp mRNA linear EST 11-DEC-2000  
 BF529174  
 LOCUS 602042051F1 NCI\_CGAP\_Brn67 Homo sapiens cDNA clone IMAGE:4179828  
 DEFINITION 5', mRNA sequence.

ACCESSION BF529174  
 VERSION BF529174.1 GI:11616455  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/  
 1 (bases 1 to 898)

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov

COMMENT Tissue Procurement: David N. Louis, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.llnl.gov  
 Plate: L14M9490 row: e column: 13  
 High quality sequence stop: 587.  
 Location/Qualifiers

FEATURES  
 source

1. 898  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_1lb="IMAGE:4179828"  
 /clone\_1lb="NCI\_CGAP\_Brn67"  
 /tissue\_type="anaplastic oligodendroglioma with 1p/19q  
 loss"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: Not 1;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2.3 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP library."

BASE COUNT 154 a 275 c 280 g 189 t

Query Match 21.0%; Score 490.8; DB 12; Length 898;

Best Local Similarity 98.1%; Pred. No. 5.8e-96; Mismatches 7; Indels 3; Gaps 2;

Matches 518; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

OY 318 GTACTGTTCTTGAGACAGCTCTCGGGGGTGAAGCTTGTGACTACTGCTAAAGAGG 377  
 |||||||  
 DB 89 GTACTGTTCTTGAGACAGCTCTCGGGGGTGAAGCTTGTGACTACTGCTAAAGAGG 148  
 |||||||

OY 378 GAGACTGAGGCGGCAAGAGGCGGAAAGTTCTTCCGACAGTTGTGTGCTGAGCACTT 437  
 |||||||  
 DB 149 GAGACTGAGGCGGCAAGAGGCGGAAAGTTCTTCCGACAGTTGTGTGCTGAGCACTT 208  
 |||||||

OY 438 CTGCGACAGTACTTCATCTGCGACAGAGACTTAAAGCCGAGAACTGCTTTGGATGA 497  
 |||||||  
 DB 209 CTGCGACAGTACTTCATCTGCGACAGAGACTTAAAGCCGAGAACTGCTTTGGATGA 268  
 |||||||

OY 498 GAAAAACAACATCCGATTCAGATTCGCGATGCGCTCCCTGAGGTGGGGAGACGCT 557  
 |||||||  
 DB 269 GAAAAACAACATCCGATTCAGATTCGCGATGCGCTCCCTGAGGTGGGGAGACGCT 328  
 |||||||

QY 558 CCGGAGACAGCTGGCGGTCGCCCATATGCGTCCAGAGGTGATTAAGGGGAAAA 617  
 DB 329 CCGGAGACAGCTGGCGGTCGCCCATATGCGTCCAGAGGTGATTAAGGGGAAAA 388  
 QY 618 ATATGATGCGCGCGGCGGCAACATGTGAGCTGTGAGTATCTCTTCCGCTCTGT 677  
 DB 389 ATATGATGCGCGCGGCGGCAACATGTGAGCTGTGAGTATCTCTTCCGCTCTGT 448  
 QY 678 GGGGGCTCTCCCTTGTATGACGACACATCCGCGAGCTGCTGGAGAGTGAAGCGGG 737  
 DB 449 GGGGGCTCTCCCTTGTATGACGACACATCCGCGAGCTGCTGGAGAGTGAAGCGGG 508  
 QY 738 CGCTTCCACATGCGCCACATTCATTCCTCCAGATTCGCGAGAGCTCTGAGGGGAAT--G 795  
 DB 509 CGCTTCCACATGCGCCACATTCATTCCTCCAGATTCGCGAGAGCTCTGAGGGGAATTTGA 568  
 QY 796 ATCGAGTGGAGCCCG-AAAAAGGCTCAGCTGTGAGCAATTCAGAA 842  
 DB 569 TCGAAGTGGAGCCCGAAAAAGGCTCAGCTGTGAGCAATTCAGAA 616

RESULT 13  
 BQ444032 600 bp mRNA linear EST 29-MAY-2002

LOCUS BQ444032  
 DEFINITION UI-M-EMO-bx1-p-15-0-UI.r1 NIH\_BMAP\_EMO Mus musculus cDNA clone  
 IMAGE:5708798 5', mRNA sequence.

ACCESSION BQ444032  
 VERSION BQ444032.1 GI:21247144

KEYWORDS EST.  
 SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 JOURNAL 1 (bases 1 to 600)  
 COMMENT NIH-MGC http://mgc.ncl.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Dr. James Lin, University of Iowa  
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)

FEATURES  
 source  
 location/Qualifiers  
 1..600

/organism="Mus musculus"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone\_lib="NIH\_BMAP\_EMO"  
 /clone\_1ib="NIH\_BMAP\_EMO"  
 /issue\_type="whole brain"  
 /dev\_stage="embryo 15.5 dpc"  
 /lab\_host="DH10B (TI phage resistant)"  
 /note="Organ: brain; Vector: pYX-Asc; Site\_1: Ecor I;  
 Site\_2: Not I; The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured mRNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was size selected according to mRNA size fraction,  
 ligated with Ecor I adaptor, digested with Not I, and then  
 cloned directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is GTGGGTGGA. This library was created for the  
 University of Iowa Mouse Brain Molecular Anatomy Project  
 (BMAP). Gene Discovery in the Developing Mouse Nervous  
 System, supported by National Institutes of Mental Health

BASE COUNT 145 a 156 c 173 g 125 t 1 others  
 ORIGIN

Query Match 21.0%; Score 490.2; DB 14; Length 600;  
 Best Local Similarity 91.4%; Pred. No. 7e-96;  
 Matches 541; Conservative 0; Mismatches 49; Indels 2; Gaps 2;

QY 274 CATGCTCCTAAGCTCCACGAGCTGTACGAGACAAAGAAATATTGTACCTGCTTGGAG 333  
 DB 11 CAGCTCTCAAGCTCCACGAGCTGTACGAGACAAAGAAATATTGTACCTGCTTGGAG 70  
 QY 334 CAGCTCTCGGGGGGTGAGCTATTGCACTACCTGTTAAAGAGGAGACTGACCCCAAG 393  
 DB 71 CAGCTTCTGTGTGTGAGCTGTTCGACTACCTGTTAAAGAGGAGACTGACCCCAAG 130  
 QY 394 GAGGCCGAAAGTCTTCCGCGCAGATTGTGTGCGTGTGAGACTTTCGACAGTACTCC 453  
 DB 131 GAGGCCGGAAGTCTTCCGCGCAGATTGTGTGCGTGTGAGACTTTCGACAGTACTCC 190  
 QY 454 ATCTGCCAGAGAGCTTAAAGCCGGAACCTGCTTTGGTGAAGAAACACATCCGC 513  
 DB 191 ATCTGTACAGAGACTTGAAGCCAGAGAACTGCTGTTGATGAGAAACACATCCGC 250  
 QY 514 ATTCGAGACTTCCGCAATGGCTCCCTGCGAGCTGAGGAGACACCTCTGAGACAGCTGC 573  
 DB 251 ATCGCAGACTTGTGTATGGGTCCCTGCAAGTGGGGAGACGCTCTGAGAGACAGCTGC 310  
 QY 574 GGGTCCCCCATTTATGCGTGTCCAGAGGTGATTAAGGGGGAATATGATGCGCCGG 633  
 DB 311 GGGTCCCCCATTTATGCGTGTCCAGAGGTGATTAAGGGGGAATATGATGCGCCGG 370  
 QY 634 GCAGACATGAGAGCTGTGAGTGCATCTCTGCGCCCTGCGTGGGGGCTGACCTTT 693  
 DB 371 GCAGACATGAGAGCTGTGAGTGCATCTCTGCGCCCTGCGTGGGGGCTGACCTTT 430  
 QY 694 GATGACAGAACCTCCGCCAGCTGTGAGAGGTGAAGCGGGGCTCTCCACATGCC 753  
 DB 431 GATGACAGAACCTCCGCCAGCTGTGAGAGGTGAAGCGGGGCTCTCCACATGCC 489  
 QY 754 CACTTATCTCTCCATGATTCGAGAGCTCTGAGGGGATGATGAGAGGAGCCCGGA 813  
 DB 490 CACTTATCTCTCCATGATTCGAGAGCTCTGAGGGGATGATGAGAGGAGCCCGGA 549  
 QY 814 AAAAGCTCAGTCTGAGAGCAATTCAGAAACATCTGTGATCTAGAGCGGA 865  
 DB 550 AAAAGCTCAGTCTGAGAGCAATTCAGAAACATCTGTGATCTAGAGCGGA 600

RESULT 14  
 BE796285 753 bp mRNA linear EST 12-JAN-2001  
 LOCUS BE796285  
 DEFINITION G02258580F1 NIH\_MGC\_85 Homo sapiens cDNA clone IMAGE:4341838 5',  
 mRNA sequence.

ACCESSION BE796285  
 VERSION BE796285.1 GI:12101339

KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 JOURNAL NIH-MGC http://mgc.ncl.nih.gov/  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Louis Staudt, M.D., Ph.D.  
 cDNA library preparation: Life Technologies, Inc.  
 cDNA library arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov  
Plate: LHM956 row: c column: 23  
High quality sequence stop: 691.

## FEATURES

## Source

1. 753  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4341838"  
/issue="1 lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lymph. Vector: pCMV-Sport6; Site: 1: NotI;  
Site: 2: SalI. Cloned unidirectionally. Oligo-dT primed.  
Average insert size 1.867 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH-MGC library."

BASE COUNT 156 a 212 c 262 g 123 t  
ORIGIN

Query Match 20.98; Score 487.8; DB 12; Length 753;  
Best Local Similarity 96.88; Pred. No. 2.5e-95;  
Matches 572; Conservative 0; Mismatches 12; Indels 7; Gaps 7;

QY 858 AGGGGGAACAGACGACGACCGCTGGAGCCAGCCCTGGCCGGGTAGCCAT 917  
DB 167 AGGGGGAACAGACGACGACCGCTGGAGCCAGCCCTGGCCGGGTAGCCAT 226  
QY 918 GCGGAGCTGCGATCAACAGAGAGCTGAGCCCGGAGCTCTAGAGAGCATCACT 977  
DB 227 GCGGAGCTGCGATCAACAGAGAGCTGAGCCCGGAGCTCTAGAGAGCATCACT 286  
QY 978 GGGCTGCTTCAGGAGCGCGAGAGCTGCGAGCTGCGAGTGAAGAGAGCA 1037  
DB 287 GGGCTGCTTCAGGAGCGCGAGAGCTGCGAGCTGCGAGTGAAGAGAGCA 346  
QY 1038 AGAAGATGATATATATCTGCTTTGGATCGGAAGAGCGGTATCCAGCTGTAGGA 1097  
DB 347 AGAAGATGATATATATCTGCTTTGGATCGGAAGAGCGGTATCCAGCTGTAGGA 406  
QY 1098 CAGAGACCTGCTCCCGGAGTATGATGACCCCGGAGAGCTGTGATTTCCAT 1157  
DB 407 CAGAGACCTGCTCCCGGAGTATGATGACCCCGGAGAGCTGTGATTTCCAT 466  
QY 1158 GCTGAGCTGACGAGGAGCGGAGCAGAGAGCTGATGAAATCTTGAATCAC 1217  
DB 467 GCTGAGCTGACGAGGAGCGGAGCAGAGAGCTGATGAAATCTTGAATCAC 525  
QY 1218 CGATGCGGGGGGTGGTGGCTCCCTGATCCACCGAGCGGGCTTGGAGATGGCCAGCA 1277  
DB 526 CGATGCGGGGGGTGGTGGCTCCCTGATCCACCGAGCGGGCTTGGAGATGGCCAGCA 585  
QY 1278 CAGCAGAGATCCCGTACGTCAGT-GGAGCTTCCAGAGGCTGTCTCCAGCCCTCTAA 1336  
DB 586 CAGCAGAGATCCCGTACGTCAGT-GGAGCTTCCAGAGGCTGTCTCCAG-CCCTCTAA 644  
QY 1337 GCAGCCCAAGAGATCCCGTCTTTCTTTTTCACCGGAGCGGGGGCTGAGATGAGGCTC 1396  
DB 645 GCAG-CCAAAGAGATCCCGTCTTTCTTTTTCACCGGAGCG-GGGGCTGAGATGAGGCTC 702  
QY 1397 GAGCGGGGGGTCCCGGCTCCAAAGAGAGCTGCTTCTGGGGGCC 1447  
DB 703 GA-CCGGGGGTCCCGG-TTCCAAAGAGAGAGCTGCTTCTGGGGGCC 751

RESULT 15 614 bp mRNA linear EST 14-MAR-2002  
BM944236  
LOCUS BM944236  
DEFINITION UT-M-EHOP-bvq-a-22-0-UT.r1 NIH-BMAP\_EHOP Mus musculus cDNA clone  
IMAGE:5695389 5, mRNA sequence.  
ACCESSION BM944236  
VERSION BM944236.1 GI:19427821  
KEYWORDS EST.  
SOURCE house mouse.

## ORGANISM

Mus musculus

## REFERENCE

1 (bases 1 to 614)

## AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pyx-5.

## FEATURES

## Source

1. 614  
/organism="Mus musculus"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5695389"  
/issue="whole brain"  
/dev\_stage="embryo 18.5 dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/note="Organ: brain; Vector: pyx-asc; Site: 1: EcoR I;  
Site: 2: Not I; The library was constructed according to  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured mRNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with an  
oligo-dT primer containing a Not I site. Double stranded  
cDNA was size selected according to mRNA size fraction,  
ligated with EcoR I adaptor, digested with Not I, and then  
cloned directionally into pyx-asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CAGCAGCAGC. This library was created for the  
University of Iowa Mouse Brain Molecular Anatomy Project  
(BMAP): 'Gene Discovery in the Developing Mouse Nervous  
System', supported by National Institute of Mental Health  
(NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT 153 a 165 c 183 g 113 t  
ORIGIN

Query Match 20.48; Score 476.2; DB 14; Length 614;  
Best Local Similarity 89.18; Pred. No. 7.6e-93;  
Matches 548; Conservative 0; Mismatches 63; Indels 4; Gaps 3;

QY 440 GCCACAGTACTCCATGTCGACAGAGACCTAAAGCCGAGAACCTGCTTTGATGAGA 499  
DB 1 GCCATAGCTATCATCTGTCACAGAGACTGAAAG-CAGAGAACTGCTGTTGATGAGA 59  
QY 500 AAAACAACATCCGATTCGACACTTGGCATGGGCTCCCTGAGGTGGGGAGACCTCC 559  
DB 60 AAAACAACATCCGATTCGACACTTGGCATGGGCTCCCTGCAAGTGGGGAGACCTCC 119  
QY 560 TGGAGACAGCTGGGGGGCCCCCATTTATGGTGCAGAGGTGATTAAGGGGAAAAAT 619  
DB 120 TGGAGACAGCTGGGGGGCCCCCATTTATGGTGCAGAGGTGATTAAGGGGAAAAAT 179  
QY 620 ATGATGGCCGCGGAGAGACATGTGAGCTGTGAGTATCTTTCGCTGCTGCTG 679  
DB 180 ATGATGGCCGCGGAGAGACATGTGAGCTGTGAGTATCTTTCGCTGCTGCTG 239  
QY 680 GGGCTGCTGCTTTGATGAGCAAACTTCGCGACCTGCTGAGAGAGTGAAGCGGGCG 739  
DB 240 GGGCTGCTGCTTTGATGAGCAAACTTCGCGACCTGCTGAGAGAGTGAAGCGGGCG 299  
QY 740 TCTTCACATCCCACTTCAATCTCCAGATTTGCCAAGAGCTCTGAGAGGAATGATCG 799

Db	300	TCCTCCACATGCGCTCACTTCATCCCTCCACAGACTGCCAGAGCGCTCTCAGAGGGATGATG	359
QY	800	AAGTGAAGCCCCGAAAAAAGAGCTAGTCTGGAGCAAAATTCCAGAAACA--TCCTTGGTACT	857
Db	360	AAGTGAAGCCCCGAAAAAGGCTAGTCTGGAGCAAAATTCCAGAAACAATCCCTTGGTATGCT	419
QY	858	AGGCGGGAAACACGAGACC--AGACCCCTGCTCTGGAGCCAGCCGCTGGCCGCGCGGTAGCA	916
Db	420	GGGCGGGAAACACGAAACCAAGACCTTGCCTGGAGCCAGCCGCGCGCGGCAAGATTAGCA	479
QY	917	TGGCGAGGCTGGCATTCACAAGGAGAGCTGGACCCCGACGTCCTAGAGACGATGGCATAC	976
Db	480	TGGCGAGGCTGGCTTCCAAATGGGAGACTGGACCTGCACGTTCTGGAAAGCATGGCGCTGC	539
QY	977	TGGGCTGCTTCAGGAGCACCGCAGAGGCGTGATCGCGAGCTGGCCGACGTAGAGAGAGACAAC	1036
Db	540	TGGGCTGCTTCAGAGACCGCGAGCGGCTGCACAGAGAACTGCCGAAGGAGAGGAAAMCC	599
QY	1037	AAGAAAGATGATAT 1051	
Db	600	AAGAAAGATGATCT 614	

Search completed: April 22, 2003, 00:23:49  
Job time : 2137.25 secs

GenCore version 5.1.4-P5-A578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:23:30 ; Search time 320.629 Seconds  
(without alignments)  
16393.314 Million cell updates/sec

Title: US-10-003-690-3  
Perfect score: 2334  
Sequence: 1 atgtctcggggcgaagga.....ccaacggagccctctgcc 2334

Scoring table: IDENTITY\_MDC  
Gap 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq.101002.\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2256	96.7	2385	22	AAS06717
2	2256	96.7	2897	24	AA026464
3	961	41.2	2025	24	ABA02995
4	961	41.2	2217	24	ABA02994
5	955.8	41.0	2647	24	AA034315
6	658.8	28.2	906	22	ABA08296
7	614	26.3	614	22	ABA08924
8	512	21.9	512	24	ABK70216
9	470.8	20.2	2720	23	ABL04689

10	264.2	11.3	2954	22	AAK94923	Human full-length
11	264.2	11.3	3269	22	AA018831	Human kinase (PKIN
12	264.2	11.3	3312	24	ABA05739	Human neuronal ser
13	264.2	11.3	3392	24	ABA05740	Human neuronal ser
14	261.6	11.2	1594	22	AA044655	Novel protein kina
15	259.4	11.1	3170	24	ABA05737	Marine neuronal se
16	259.4	11.1	3250	24	ABA05738	Marine neuronal se
17	257.8	11.0	316	22	ABA13768	Human neuronal se
18	257.8	11.0	316	22	AA027156	Human neuronal se
19	257.8	11.0	316	22	ABA03545	Human neuronal se
20	240.6	10.3	2112	24	AA030397	DNA encoding nove
21	240.6	10.3	2222	24	AA030398	Human PAR-1B beta
22	240.6	10.3	2715	22	AAK51482	Human polynucleoti
23	238.8	10.2	2663	23	ABL05773	Drosophila melanog
24	237.8	10.2	2979	22	AAK52466	Human polynucleoti
25	229	9.8	4354	23	ABL07437	Drosophila melanog
26	221.4	9.5	2352	21	AAK68823	Human protein kina
27	221.4	9.5	2533	21	AAK82952	Human keratinocyte
28	221.4	9.5	2632	21	AAK82951	Human keratinocyte
29	221.4	9.5	2668	22	AAK68822	Human keratinocyte
30	221.4	9.5	4699	22	AAH76213	Human kinase PKIN-
31	212.8	9.1	1536	14	AA047789	SHP gene. Yeast
32	212.4	9.1	2222	24	AA033068	Human PAR-1B beta
33	211.2	9.0	2118	23	ABL11605	Drosophila melanog
34	211.2	9.0	4118	23	ABL11604	Drosophila melanog
35	207	8.9	2766	23	ABL26615	Drosophila melanog
36	207	8.9	3957	21	AAK77495	Human ORFX ORF3050
37	207	8.9	4321	22	AA044652	Novel protein kina
38	207	8.9	4381	24	AA034304	Human PKIN-7 cDNA.
39	206.6	8.9	1647	18	AA085925	Mammalian AMPK alp
40	204	8.7	1349	22	AA199637	Human expressed po
41	204	8.7	1349	22	ABK43708	DNA encoding novel
42	202.4	8.7	2361	22	AAS06718	Polynucleotide seq
43	201	8.6	859	22	AAK91635	Human cDNA 5'-end
44	201	8.6	859	22	AAK93561	Human cDNA clone r
45	201	8.6	2024	22	AAK94157	Human full-length

## ALIGNMENTS

RESULT 1  
ID AAS06717 standard; cDNA: 2385 BP.  
AC AAS06717;  
XX  
XX  
DT 12-SEP-2001 (first entry)  
XX  
XX Polynucleotide sequence encoding human protein kinase #17.  
DE  
XX  
XX Human; protein kinase; PKR; GSK; cancer; cardiovascular disease;  
KW metabolic disorder; immune related disease; neurological disorder;  
KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
KW reproductive disorder; gene therapy; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200138503-A2.  
XX  
XX  
XX 31-MAY-2001.  
XX  
XX 22-NOV-2000; 2000MO-US32085.  
XX  
XX 24-NOV-1999; 99US-0167482.  
XX  
XX (SUGEN-) SUGEN INC.  
XX  
XX Plozman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
PI Pliaganan P, Clary D;  
XX  
XX WPI: 2001-343950/36.  
DR  
DR P-PSDB; AA003517.

XX Nucleic acids encoding human kinase polypeptides, useful for preventing  
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
PT neuronal-associated diseases, and microbial infections -

PS Example 1: Figure 1: 433pp; English.

CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The  
CC novel protein kinases have been identified as members of the tyrosine  
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides  
CC encoding protein kinases and the polypeptides may be used in the  
CC prevention, diagnosis and treatment of diseases associated with  
CC inappropriate kinase expression. For example, they may be used to treat  
CC cancers (especially cancers of hematopoietic origin), cardiovascular  
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
CC immune related diseases (e.g. rheumatoid arthritis), neurological  
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
CC Additionally, polynucleotides encoding protein kinases may be  
CC used for gene therapy and as DNA probes in diagnostic assays.  
CC The protein kinase polypeptides may be used as antigens in the production  
CC of antibodies against the protein kinases and in assays to identify  
CC modulators of protein kinase expression and activity.

XX Sequence 2385 BP; 494 A; 768 C; 731 G; 392 T; 0 other;

Query Match 96.7%; Score 2256; DB 22; Length 2385;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 CAGGCCCAATATGTGGCCCTATGCTGGTGAAGAGCCTGGGCAAGAGACAGACAGG 138  
DB 127 CACCCCAATATGTGGCCCTATGCTGGTGAAGAGCCTGGGCAAGAGACAGACAGG 186  
QY 139 CTGTTAACTCGGGGTCCATGTCATACGAGGCTGCAAGAGTCCATCAAGATCGGAC 198  
DB 187 CTGTTAACTCGGGGTCCATGTCATACGAGGCTGCAAGAGTCCATCAAGATCGGAC 246  
QY 199 CGGAGAAAGTGTGGAGTGTGCTGATGAAGTGAAGAGAGAGAGTCCCATCTGAAG 258  
DB 247 CGGAGAAAGTGTGGAGTGTGCTGATGAAGTGAAGAGAGAGAGTCCCATCTGAAG 306  
QY 259 CTCATCGAACCCCAATGCTCTCAAGTCCACAGAGCTGTACAGAACAAAGAAATTTG 318  
DB 307 CTCATCGAACCCCAATGCTCTCAAGTCCACAGAGCTGTACAGAACAAAGAAATTTG 366  
QY 319 TACCTGGTTCGAGACAGTCTCGGGGGGTGAGTATTCGACTACCTGGTAAAGAGGG 378  
DB 367 TACCTGGTTCGAGACAGTCTCGGGGGGTGAGTATTCGACTACCTGGTAAAGAGGG 426  
QY 379 AGACTGAGCCCAAGAGAGCCCGAAGTCTTCCGACAGATTGTGCTCGCTGAGACTTC 438  
DB 427 AGACTGAGCCCAAGAGAGCCCGAAGTCTTCCGACAGATTGTGCTCGCTGAGACTTC 486  
QY 439 TGCACAGACTACTCATATGTCACAGAGACTTAAAGCCGGAAGACTGCTTTTGGATGAG 498  
DB 487 TGCACAGACTACTCATATGTCACAGAGACTTAAAGCCGGAAGACTGCTTTTGGATGAG 546  
QY 499 AAAAACAACATCCGATTCGAGACTTGGCATGGGCTCCCTGAGGTGGGGGACACCTTC 558  
DB 547 AAAAACAACATCCGATTCGAGACTTGGCATGGGCTCCCTGAGGTGGGGGACACCTTC 606  
QY 559 CTGAGAGACACTGCGGCTCCCATTTATGCTGTCCAGAGAGTGTAAAGGGGAAAAA 618  
DB 607 CTGAGAGACACTGCGGCTCCCATTTATGCTGTCCAGAGAGTGTAAAGGGGAAAAA 666  
QY 619 TATATGAGCCCGGAGAGACATGTGAGCTGTGAGTATCTTTCGCTGCTGCTG 678  
DB 667 TATATGAGCCCGGAGAGACATGTGAGCTGTGAGTATCTTTCGCTGCTGCTG 726  
QY 679 GGGGCTGCGCTTTGATGAGAGCAACTCCGCGAGCTGCTGGAAGAGTGAACGGGGC 738

DB 727 GGGGCTGCGCTTTGATGAGACAACTCCGCGAGCTGCTGAGAGAGTGAACGGGGC 786  
QY 739 GTCTTCCACATGCCCACTTATTCCTCCAGATTGCCAGAGGCTCTGAGGGAATGATC 798  
DB 787 GTCTTCCACATGCCCACTTATTCCTCCAGATTGCCAGAGGCTCTGAGGGAATGATC 846  
QY 799 GAACTGAGCCCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTTGTACTTA 858  
DB 847 GAACTGAGCCCGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTTGTACTTA 906  
QY 859 GGGGGGAAACAGAGCCAGACCCCGGCTGAGAGCCAGCCCGCTGGGCGGCTAGCCATG 918  
DB 907 GGGGGGAAACAGAGCCAGACCCCGGCTGAGAGCCAGCCCGCTGGGCGGCTAGCCATG 966  
QY 919 CGGAGCTGCTCATCAACAGAGAGCTGAGACCCGAGCTGAGAGAGTGCATCAGC 978  
DB 967 CGGAGCTGCTCATCAACAGAGAGCTGAGACCCGAGCTGAGAGAGTGCATCAGC 1026  
QY 979 GGGTCTTTCAGGAGACCGGAGAGGCTGATGCGAGCTGCGCATGAGAGAGAACCA 1038  
DB 1027 GGGTCTTTCAGGAGACCGGAGAGGCTGATGCGAGCTGCGCATGAGAGAGAACCA 1086  
QY 1039 GAAAGATGATATATATATATCTTTTGGATGAGAGAGAGAGAGAGAGAGAGAG 1098  
DB 1087 GAAAGATGATATATATATATCTTTTGGATGAGAGAGAGAGAGAGAGAGAGAG 1146  
QY 1099 CAGGACCTGCTCCCGGAAATGATGTGACCCCGGAGAGCTGAGATTCCTCCATG 1158  
DB 1147 CAGGACCTGCTCCCGGAAATGATGTGACCCCGGAGAGCTGAGATTCCTCCATG 1206  
QY 1159 CTGAGCCGTCAAGGAGAGAGGCGGACAGAGCGGAAGTCCATGAAATCCTGAGATCAC 1218  
DB 1207 CTGAGCCGTCAAGGAGAGAGGCGGACAGAGCGGAAGTCCATGAAATCCTGAGATCAC 1266  
QY 1219 GATGCGGGGGGTGGTGGTCCCTGCTTACCAACCGGAGGCTTGGAGATGGCCACGAC 1278  
DB 1267 GATGCGGGGGGTGGTGGTCCCTGCTTACCAACCGGAGGCTTGGAGATGGCCACGAC 1326  
QY 1279 AGCCAGAGATCCCGTAGCTGAGTGAAGAGCTGACAGGCTGCTCCAGCCCTTAAGC 1338  
DB 1327 AGCCAGAGATCCCGTAGCTGAGTGAAGAGCTGACAGGCTGCTCCAGCCCTTAAGC 1386  
QY 1339 AGCCCAAGAGATCCCGTCTTTTCTTTTACCGGAGCGGGGGCTGAGAGAGCTCGA 1398  
DB 1387 AGCCCAAGAGATCCCGTCTTTTCTTTTACCGGAGCGGGGGCTGAGAGAGCTCGA 1446  
QY 1399 GGGGGGGGCTCCCGAGCTTCAAAAAGCAGAGAGCTGCTTCCGGGGGCCCCAGGGGTGG 1458  
DB 1447 GGGGGGGGCTCCCGAGCTTCAAAAAGCAGAGAGCTGCTTCCGGGGGCCCCAGGGGTGG 1506  
QY 1459 GGGCGCGGGGAGACCCCGCCCGCCAGTGCAGCTCCGCTCCACACCCCTCCGCGCCCA 1518  
DB 1507 GGGCGCGGGGAGACCCCGCCCGCCAGTGCAGCTCCGCTCCACACCCCTCCGCGCCCA 1566  
QY 1519 GGGTCCCGGCTCTCTGCTGGGAGACCCCTTGCACCTGCTTGCACACGCCCCGGGC 1578  
DB 1567 GGGTCCCGGCTCTCTGCTGGGAGACCCCTTGCACCTGCTTGCACACGCCCCGGGC 1626  
QY 1579 AGTCCACCCGGGAGCCCGGGGAGCAAAACACACCCCGGAGCGGGGGTGGGGGA 1638  
DB 1627 AGTCCACCCGGGAGCCCGGGGAGCAAAACACACCCCGGAGCGGGGGTGGGGGA 1686  
QY 1639 GCGGCTGAGAGAGTGTCTTCACTCACTCCGCAACAGACTTCTGGGCTCCCTGCTT 1698  
DB 1687 GCGGCTGAGAGAGTGTCTTCACTCACTCCGCAACAGACTTCTGGGCTCCCTGCTT 1746  
QY 1699 CACCGGCGAAGATGACAGTCCCTACCGCTGAGAGATGCTCAAGCTTCAAGAGTCC 1758  
DB 1747 CACCGGCGAAGATGACAGTCCCTACCGCTGAGAGATGCTCAAGCTTCAAGAGTCC 1806  
QY 1759 TCCCGGAGCTGGCAAAAGCTCTCTGCTTGGGAACTTCACTCTCTTGAACAAAGAA 1818  
DB 1807 TCCCGGAGCTGGCAAAAGCTCTCTGCTTGGGAACTTCACTCTCTTGAACAAAGAA 1866

QY 1819 CAATATTCCTGCTGCTAAAGGACAAACCTCTCAGCAGCATCAAGACAGATCTCCAT 1878  
 DB 1867 CAAATATTCCTGCTGCTAAAGGACAAACCTCTCAGCAGCATCAAGACAGATCTCCAT 1926  
 QY 1879 GCCTTCTGCTGATCCCGAGCTGATGATGCTGTCTCAGACAGCAGCTTCAGAGCC 1938  
 DB 1927 GCCTTCTGCTGATCCCGAGCTGATGATGCTGTCTCAGACAGCAGCTTCAGAGCC 1986  
 QY 1939 GAGTACAAAGGACAGTGGGCGCCCTCCCTCTTCAAAAACCCGCTCCGCTTCAGAGTGAC 1998  
 DB 1987 GAGTACAAAGGACAGTGGGCGCCCTCCCTCTTCAAAAACCCGCTCCGCTTCAGAGTGAC 2046  
 QY 1999 ATCAGCTCTCTGAGAGTCCAGAGCTCTCCCGCAGAGGAGGAGGAGGAGTGGTGGC 2058  
 DB 2047 ATCAGCTCTCTGAGAGTCCAGAGCTCTCCCGCAGAGGAGGAGGAGGAGTGGTGGC 2106  
 QY 2059 ATCTACTCCGCTCAGCTTCTACTCTCATCTGCGGGTCCAGCGCTGGTTCAGAGGAGTGGT 2118  
 DB 2107 ATCTACTCCGCTCAGCTTCTACTCTCATCTGCGGGTCCAGCGCTGGTTCAGAGGAGTGGT 2166  
 QY 2119 GAGACCATCCAGGACAGCTCTCTGAGCAGCTCATGACCAAGCCCTCCGCTGAGGCGCTTGCA 2178  
 DB 2167 GAGACCATCCAGGACAGCTCTCTGAGCAGCTCATGACCAAGCCCTCCGCTGAGGCGCTTGCA 2226  
 QY 2179 GAGCAGAAAGAACGGGGGCCCAAGACCCGCTGCTGCTGCTCCCAACCCGAAAGCTGACGCC 2238  
 DB 2227 GAGCAGAAAGAACGGGGGCCCAAGACCCGCTGCTGCTGCTCCCAACCCGAAAGCTGACGCC 2286  
 QY 2239 CCACCCGGGCGCCGACAGACCCAGAGCTGAGACAGCTCTCCCGCGAGGAGGAGGAGGAGGAGGAG 2298  
 DB 2287 CCACCCGGGCGCCGACAGACCCAGAGCTGAGACAGCTCTCCCGCGAGGAGGAGGAGGAGGAGGAG 2346  
 QY 2299 AAGAAGCTCTCTGCGCCACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2334  
 DB 2347 AAGAAGCTCTCTGCGCCACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2382  
 RESULT 2  
 AAD26464 ID AAD26464 standard; cDNA; 2897 BP.  
 AC AAD26464;  
 DT 26-MAR-2002 (first entry)  
 DE Human kinase PKIN-17 cDNA.  
 XX  
 KW Human: kinase; PKIN-17; cancer; leukemia; adenocarcinoma; osteoporosis;  
 KW Immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;  
 KW Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anemia;  
 KW allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;  
 KW autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;  
 KW Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;  
 KW Rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;  
 KW hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;  
 KW cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;  
 KW congestive heart failure; ischemic heart disease; lung tumour; gout;  
 KW fatty liver; Niemann-Pick's disease; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..2385  
 FT /\*tag= a  
 FT /product= "Human PKIN-17 protein"  
 XX  
 PN W0200196547-A2.  
 XX  
 PD 20-DEC-2001.  
 XX  
 PF 14-JUN-2001; 2001W0-US194444.

PR 15-JUN-2000; 2000US-212073P.  
 PR 23-JUN-2000; 2000US-213467P.  
 PR 30-JUN-2000; 2000US-215651P.  
 PR 07-JUL-2000; 2000US-216605P.  
 PR 13-JUL-2000; 2000US-218372P.  
 PR 25-AUG-2000; 2000US-228056P.  
 PA (INCYTE GENOMICS INC.  
 XX  
 PI Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;  
 PI Gandhi AR, Tripathy CM, Walla NK, Yeo MG, Lu DM, Greenwald SR;  
 PI Ramkumar J, Griffin JR, Kearney L, Burford N, Nguyen DB, Tang YT;  
 PI Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;  
 PI Lo TP, Khan F, Reardon SA, Azimzal Y, Policky JL, Ding L;  
 PI Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;  
 DR WPI: 2002-090207/12.  
 DR P-PSDB; AAE16271.  
 PT New polypeptides, useful for diagnosing, treating or preventing  
 PT disorders of growth and development, cardiovascular and lipid, and  
 PT diseases such as cancer, comprise human kinase polypeptides  
 PS Claim 5; Page 188-189; 197pp; English.  
 XX  
 CC The invention relates to human kinase PKIN proteins and their  
 CC corresponding cDNAs. A composition containing PKIN agonist is useful for  
 CC treating a disease or condition associated with decreased expression of  
 CC PKIN and a composition comprising PKIN antagonist is useful for treating  
 CC a disease or condition associated with overexpression of PKIN. The  
 CC disorders include cancer (leukemia, adenocarcinoma, lymphoma, melanoma,  
 CC myeloma, sarcoma, teratocarcinoma, Hodgkin's disease), immune disorder  
 CC (acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,  
 CC atherosclerosis, anemia, allergies, adult respiratory distress syndrome,  
 CC autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes  
 CC mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,  
 CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,  
 CC rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,  
 CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)  
 CC growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,  
 CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts), cardio  
 CC vascular disease (arteriovenous fistula, hypertension, vasculitis,  
 CC aneurysms, congestive heart failure, angina pectoris, myocarditis,  
 CC ischemic heart disease, chronic bronchitis, lung tumours), lipid  
 CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,  
 CC hypercholesterolemia, obesity). PKIN DNA is useful for assessing  
 CC toxicity of a test compound and in gene therapy. The present sequence  
 CC is human PKIN-17 cDNA.  
 XX  
 SQ Sequence 2897 BP; 599 A; 921 C; 877 G; 500 T; 0 other;  
 Query Match 96.7%; Score 2256; DB 24; Length 2897;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 79 CAGGCCCAATATGTTGGGCCCCATATGCGGTGAGAAAGACGCTGGCAAGACAGACAGG 138  
 DB 127 CAGGCCCAATATGTTGGGCCCCATATGCGGTGAGAAAGACGCTGGCAAGACAGACAGG 186  
 QY 139 CTGTTTAACCTCGGGGTCACATGATCAGAGGTCAGAAAGTGGCCATCAAGATCTGAAAC 198  
 DB 187 CTGTTTAACCTCGGGGTCACATGATCAGAGGTCAGAAAGTGGCCATCAAGATCTGAAAC 246  
 QY 199 CGGGAAGAGCTGTGAGAGTGGTGTGATGAGAGTGGAGCGGAGATGCCATCTCTGAAG 258  
 DB 247 CGGGAAGAGCTGTGAGAGTGGTGTGATGAGAGTGGAGCGGAGATGCCATCTCTGAAG 306  
 QY 259 CTCATGGAACCAACCATGCTCTCAAGCTCCACGAGCTCTACGAGAACAATAATTTG 318  
 DB 307 CTCATGGAACCAACCATGCTCTCAAGCTCCACGAGCTCTACGAGAACAATAATTTG 366  
 QY 319 TACCTGTTTGGAGCAGAGTCTCGGGGGTGAAGCTATTTGACTTACTGTTAAAGAGGG 378

Db 367 TACCTGGTTGAGACACGCTCGGGGGGTGAGTATTCGACTACCTGGTAAAGAGGG 426  
 QY 379 AGACTGAGCCCAAGAGAGAGCCGGAAGTTCTCCGACAGATTGTGTCGCGTGACATTC 438  
 Db 427 AGACTGAGCCCAAGAGAGAGCCGGAAGTTCTCCGACAGATTGTGTCGCGTGACATTC 486  
 QY 439 TGGCAGACACTACTCATCTGCGACAGAGACCTAAAGCCCGAGAACCTGCTTTTGATGAG 498  
 Db 487 TGGCAGACACTACTCATCTGCGACAGAGACCTAAAGCCCGAGAACCTGCTTTTGATGAG 546  
 QY 499 AAAAACAACATCCGATGTCAGACTTCGGCATGGGCTCCCTCAGATGGGGAGACACCTTC 558  
 Db 547 AAAAACAACATCCGATGTCAGACTTCGGCATGGGCTCCCTCAGATGGGGAGACACCTTC 606  
 QY 559 CTGAGAGACACACTGCGGGTCCCGCATTCGCTGTCAGAGAGTGATTAAGGGGAGAAAA 618  
 Db 607 CTGAGAGACACACTGCGGGTCCCGCATTCGCTGTCAGAGAGTGATTAAGGGGAGAAAA 666  
 QY 619 TATGATGGCCCGCGGAGACATGTGAGCTGTGAGTCACTCTTCGCGCTGCTGATG 678  
 Db 667 TATGATGGCCCGCGGAGACATGTGAGCTGTGAGTCACTCTTCGCGCTGCTGATG 726  
 QY 679 GGGGCTCGCCCTTTGATGAGAGACAACTCCCGCAGCTGCTGGAAGAGGTGAACGGGGC 738  
 Db 727 GGGGCTCGCCCTTTGATGAGAGACAACTCCCGCAGCTGCTGGAAGAGGTGAACGGGGC 786  
 QY 739 GTCTTCACATGCCCCACTTCATTCCTCAGATTGCCAGAGCTCCTCGAGGGGAATGATC 798  
 Db 787 GTCTTCACATGCCCCACTTCATTCCTCAGATTGCCAGAGCTCCTCGAGGGGAATGATC 846  
 QY 799 GAAATGAGCCCGGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTTGTGTAACCTA 858  
 Db 847 GAAATGAGCCCGGAAAAAGGCTCAGTCTGAGCAAAATTCAGAAACATCTTGTGTAACCTA 906  
 QY 859 GGGGGAAGAACGAGAGCCGAGACCCGTCGAGAGAGAGCCCGGGGGGAGGAGCATG 918  
 Db 907 GGGGGAAGAACGAGAGCCGAGACCCGTCGAGAGAGAGCCCGGGGGGAGGAGCATG 966  
 QY 919 CGAGAGCTGCCATCCACGAGAGAGTGAACCCGACGTCCTAGAGAGCATGTCATCTG 978  
 Db 967 CGAGAGCTGCCATCCACGAGAGAGTGAACCCGACGTCCTAGAGAGCATGTCATCTG 1026  
 QY 979 GGGCTCTTCAGAGAGCCGGAAGAGTGCATGCGAGCTGCGAGTGAAGAGAGAGAACCAA 1038  
 Db 1027 GGGCTCTTCAGAGAGCCGGAAGAGTGCATGCGAGCTGCGAGTGAAGAGAGAGAACCAA 1086  
 QY 1039 GAAAAGATGATATATATATCTGCTTTTGATCGAGAGAGAGCGGTATCCCGATGAGAGAC 1098  
 Db 1087 GAAAAGATGATATATATATCTGCTTTTGATCGAGAGAGAGCGGTATCCCGATGAGAGAC 1146  
 QY 1099 CAGGAGCTGCCCTCCCGGAATGATGTTGACCCCGGAAAGCGTGTGATTCCTCCATG 1158  
 Db 1147 CAGGAGCTGCCCTCCCGGAATGATGTTGACCCCGGAAAGCGTGTGATTCCTCCATG 1206  
 QY 1159 CTGAGAGCTCAGCGGAAAGCGGAGACAGAGGGAAGTCCATGGAAGTCTCTGAGCATCAC 1218  
 Db 1207 CTGAGAGCTCAGCGGAAAGCGGAGACAGAGGGAAGTCCATGGAAGTCTCTGAGCATCAC 1266  
 QY 1219 GATGCGGGGGTGTGCTGCTGTCACCAACCCGAGCGGCTTGGAGATGGCCCAAGCAC 1278  
 Db 1267 GATGCGGGGGTGTGCTGCTGTCACCAACCCGAGCGGCTTGGAGATGGCCCAAGCAC 1326  
 QY 1279 AGCCAGAGATCCCTAGACGTGAGAGCTTCACAGGGTGTCTCTCCAGCCCTCTCAAC 1338  
 Db 1327 AGCCAGAGATCCCTAGACGTGAGAGCTTCACAGGGTGTCTCTCCAGCCCTCTCAAC 1386  
 QY 1339 AGCCAGAGATCCCTAGACGTGAGAGCTTCACAGGGTGTCTCTCCAGCCCTCTCAAC 1398  
 Db 1387 AGCCAGAGATCCCTAGACGTGAGAGCTTCACAGGGTGTCTCTCCAGCCCTCTCAAC 1446  
 QY 1399 GGGGAGGGGCTCCCGAGACTTCAAAAAGCAGAGGCTTCCTCGGGGGCCCGAGGGGTGGG 1458  
 Db 1447 GGGGAGGGGCTCCCGAGACTTCAAAAAGCAGAGGCTTCCTCGGGGGCCCGAGGGGTGGG 1506

QY 1459 GGGCGGGGAGAGACCCCGCCCGCCAGTGGCGGCTCCACACCCCTGGCCGGCCCCCA 1518  
 Db 1507 GGGCGGGGAGAGACCCCGCCCGCCAGTGGCGGCTCCACACCCCTGGCCGGCCCCCA 1566  
 QY 1519 GGGTCCCGGGCTCTCTGCGGGAGACCCCTTGACACTGCGCTCTGACAGCCCGGGGCC 1578  
 Db 1567 GGGTCCCGGGCTCTCTGCGGGAGACCCCTTGACACTGCGCTCTGACAGCCCGGGGCC 1626  
 QY 1579 AGTCCACCCGGGAGACCCCGGGAGAACACACACCCCGAGCCCGGGGTGGCGGGA 1638  
 Db 1627 AGTCCACCCGGGAGACCCCGGGAGAACACACACCCCGAGCCCGGGGTGGCGGGA 1686  
 QY 1639 GCGGCTGAGAGAGTCTGCTCACTCCATCCGCAACACACTTCCTGGGTCCCGCTTT 1698  
 Db 1687 GCGGCTGAGAGAGTCTGCTCACTCCATCCGCAACACACTTCCTGGGTCCCGCTTT 1746  
 QY 1699 CACCGGGCAAGATGACAGTCCCTACCGCTGAGAGATGTCAGCTTACGCGCAGAGTCC 1758  
 Db 1747 CACCGGGCAAGATGACAGTCCCTACCGCTGAGAGATGTCAGCTTACGCGCAGAGTCC 1806  
 QY 1759 TCCCGGAGCTGGGCAAAAGCTCCTGCTGCTGGGAACTTATCTCTTGGACAAAGAGA 1818  
 Db 1807 TCCCGGAGCTGGGCAAAAGCTCCTGCTGCTGGGAACTTATCTCTTGGACAAAGAGA 1866  
 QY 1819 CAATATTCCTCTGCTAAAGAGCAAACTCTCAGACAGATCAAAAGAGATGCTCCAT 1878  
 Db 1867 CAATATTCCTCTGCTAAAGAGCAAACTCTCAGACAGATCAAAAGAGATGCTCCAT 1926  
 QY 1879 GCGTTTGTGTGATCCCGAGCTGATCAGATGCTGTGTGACAGACAGCTTCAGGGCC 1938  
 Db 1927 GCGTTTGTGTGATCCCGAGCTGATCAGATGCTGTGTGACAGACAGCTTCAGGGCC 1986  
 QY 1939 GAGTACAAAGGCAAGTGGCGGCGCCCTCGCTTCCAAAAAGCCCGCTTCAGGTGAGC 1998  
 Db 1987 GAGTACAAAGGCAAGTGGCGGCGCCCTCGCTTCCAAAAAGCCCGCTTCAGGTGAGC 2046  
 QY 1999 ATCAGCTCTCTGAGGGTTCAGAGACCCCTCCCGGAGCGGAGCGGAGAGTGTGAGC 2058  
 Db 2047 ATCAGCTCTCTGAGGGTTCAGAGACCCCTCCCGGAGCGGAGCGGAGAGTGTGAGC 2106  
 QY 2059 ATCTACTCCGTCACCTTCTCTCTATCTCGGGTCCAGCCGCTGGTTCAAGGAGTGTG 2118  
 Db 2107 ATCTACTCCGTCACCTTCTCTCTATCTCGGGTCCAGCCGCTGGTTCAAGGAGTGTG 2166  
 QY 2119 GAGACATCCAGGACAGCTCTGAGACATGACACCGCCCTCGAGGAGCCCTGGCA 2178  
 Db 2167 GAGACATCCAGGACAGCTCTGAGACATGACACCGCCCTCGAGGAGCCCTGGCA 2226  
 QY 2179 GAGGAGAAAGAGGGGCGCCAGAGCCCGGCTGCTGTGCGCCACCCGAAAGCTGAGGCC 2238  
 Db 2227 GAGGAGAAAGAGGGGCGCCAGAGCCCGGCTGCTGTGCGCCACCCGAAAGCTGAGGCC 2286  
 QY 2239 CCACCGGGCGCCCAAGACCAAGAGCTGAGACGCTCTCCCGCGAGAGCCCGCCCAAGAGC 2298  
 Db 2287 CCACCGGGCGCCCAAGACCAAGAGCTGAGACGCTCTCCCGCGAGAGCCCGCCCAAGAGC 2346  
 QY 2299 AAGAAGCTCTGGGCAACCAAGGAGCCCTGTGCC 2334  
 Db 2347 AAGAAGCTCTGGGCAACCAAGGAGCCCTGTGCC 2382

RESULT 3  
 ABA02995  
 ID ABA02995 standard; cDNA; 2025 BP.

XX ABA02995;

XX 19-FEB-2002 (first entry)

XX Human protein kinase 2246 coding sequence SEQ ID NO. 3.

XX Human; protein kinase 2246; cytosolic; immunomodulator; carcinoma;

KW anti-inflammatory; analgesic; cardiovascular; cancer; sarcoma;  
 KW cellular proliferation disorder; cellular differentiation disorder;  
 KW metastatic; haematopoietic disorder; leukaemia; immune disorder;  
 KW inflammatory disorder; arthritis; autoimmune disease; diabetes mellitus;  
 KW psoriasis; Crohn's disease; cardiovascular disease; virus; pain;  
 KW gene therapy; ss.  
 OS Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT 1..2025  
 FT CDS /tag- a  
 FT /product- "protein kinase 2246"  
 PN MO200181588-A2.  
 XX  
 XX 01-NOV-2001.  
 XX  
 XX 25-APR-2001; 2001MO-US13784.  
 XX  
 XX 25-APR-2000; 2000US-199391P.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX  
 XX  
 PI Meyers R;  
 DR WPI: 2002-049281/06.  
 DR P-PSDB; AAAA7830.  
 XX  
 PT New protein kinase nucleic acid and polypeptide molecules, designated  
 PT 2246, useful for diagnosing, preventing or treating cancer or a  
 PT cellular proliferation/differentiation disorders, e.g. carcinoma,  
 PT sarcoma or leukemias -  
 XX  
 PS Claim 1; Fig 1; 11pp; English.  
 CC The invention relates to the human protein kinase 2246 gene and the  
 CC the isolated encoded polypeptide with cytosolic, immunomodulator,  
 CC anti-inflammatory, analgesic and cardiovascular activity. The 2246  
 CC nucleic acid and polypeptide are useful for diagnosing, preventing or  
 CC treating a subject having cancer or a cellular proliferation and/or  
 CC differentiation disorder or at risk of developing cancer or a cellular  
 CC proliferation and/or differentiation disorder. In particular, the  
 CC disorder includes carcinoma, sarcoma, metastatic or haematopoietic  
 CC disorders (e.g. leukemias) or cancers of the lung, breast, thyroid, head  
 CC neck, prostate or genito-urinary tract. The 2246 nucleic acid and  
 CC polypeptide are also useful for treating immune disorders, e.g.  
 CC inflammatory (e.g. respiratory inflammation or arthritis), autoimmune  
 CC disease (e.g. diabetes mellitus, psoriasis, Wegener's granulomatosis,  
 CC Crohn's disease or Grave's disease), for treating cardiovascular  
 CC diseases, endothelial cell disorder, viral diseases or pain. The nucleic  
 CC acid and polypeptide are also useful for evaluating the efficacy of a  
 CC treatment of cancer or a cellular proliferation and/or differentiation  
 CC disorder. The nucleic acid is useful for gene therapy. The present  
 CC sequence is that of the 2246 coding sequence.  
 CC  
 CC  
 XX  
 XX Sequence 2025 BP; 442 A; 641 C; 616 G; 326 T; 0 other;  
 SQ  
 Query Match 41.28; Score 961; DB 24; Length 2025;  
 Best Local Similarity 70.18; Pred. No. 1.1e-199;  
 Matches 1477; Conservative 0; Mismatches 435; Indels 195; Gaps 6;

QY 254 TGAAGTCATCGAACACCCACATGCTCTCAAGTCACAGAGCTGTACGAGACAAGAAAT 313  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 209 TGAAGTCATGAGACACCCACAGCTCTAAAGTCAGACAGAGCTTTAGAAAACAAAAT 268  
 QY 314 ATTTGTAAGTGTGTTCTGAGACAGCTCTGAGGGGGGTGAGCTATTGCACTACCTGTAAGA 373  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 269 ATTTGTAAGTGTGTTCTGAGACAGCTCTGAGGGGGGTGAGCTATTGCACTACCTGTAAGA 328  
 QY 374 AGGGGAGACTGAGCCCAAGAGGGCCGAAAGTCTTCCGCAATGTGTCTGCGCGG 433  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 329 AGGGGAGACTGAGCCCAAGAGGGCCGAAAGTCTTCCGCAATGTGTGTGCGCGG 388  
 QY 434 ACTTTCGCCACAGCTATCCATCTGACAGAGAGACTAAAGCCGAGAACTGCTTTGG 493  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 389 ACTTTCGCCACAGCTATCCATCTGACAGAGAGACTAAAGCCGAGAACTGCTTTGG 448  
 QY 494 ATGAGAAAAACAACATCCGATGTCAGACTTGGCATGGCTCCCTGCAAGTGGGGACA 553  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 449 ACAGAAAGAAACAACATCCGATGTCAGACTTGGCATGGCTCCCTGCAAGTGGGGACA 508  
 QY 554 GCCTCTGAGAGACAGCTGCGGGGTCCCTCATTTAGCTGTCCAGAGGTGATTAAGGGG 613  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 509 GCCTGTGAGAGACAGCTGCGGGGTCCCTCATTTAGCTGTCCAGAGGTGATCCGGGGG 568  
 QY 614 AAAAATATGATGGCCCGGCGAGACATGTGAGACTGTGAGATCTCTTCCGCTGC 673  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 569 AGAAGTATGAGCCCGGAGAGCGGAGCTGTGAGACTGTGAGGCTGTCTGCTTGC 628  
 QY 674 TCGTGGGGCTGCTCCCTTTGATGAGACAACCTCCGACCTGCTGAGAGAGTGAAC 733  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 629 TCGTGGGGCTGCTCCCTTTGATGAGACAACCTCCGACCTGCTGAGAGAGTGAAC 688  
 QY 734 GGGGCGCTTCCACATGCCCATCTTCAATCTCCAGATGTCAGAGCTCTGAGAGGAA 793  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 689 GGGGCGCTTCCACATGCCCATCTTCAATCTCCAGATGTCAGAGCTCTGAGAGGAA 748  
 QY 794 TGTGCAAGTGAAGCCGAAAAAAGGCTCACTGTGAGACAATTCAGAACATCTTGT 853  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 749 TGAAGCGAGTGAAGCGGAGCGGAGCTGACGTAGACATTCAGAACATATGAT 808  
 QY 854 ACCTAGCGGGGAAACAGAGCGCAACCGGTGCTGAGAGCCGCTTGGCGCGGGGTAG 913  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 809 ATTTAGGGGCAAAAGAGCCGCAACCC-----AGACCAACCCATCTTCCGCAAGTGC 862  
 QY 914 CATTGCGGAGCTGCTCAATCCAAAGAGTGAAGCCGAGCTGCTTAGAGACATGATG 973  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 863 AGATCCGCTGCTGCTCCAGCTGAGAGACATGACCCGAGCTGTGAGACATGCTACT 922  
 QY 974 CACTGGGCTGCTTCAAGGAGCGGAGAGCTGATCCGAGCTGCGAGTGAAGAGAGA 1033  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 923 CACTGGGCTGCTTCAAGGAGCGGAGAGCTGATCCGAGCTGCTGCGAGAGAGAGA 982  
 QY 1034 ACCAAGAAAAGATATATATATCTGTTTGTGATCGAAGAGAGGATATCCACTGTG 1093  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 983 ACCAAGAAAAGATATATATCTGTTTGTGATCGAAGAGAGGATATCCACTGTG 1042  
 QY 1094 AGGACACAGACTGCTCCCGGAGATGTTGACCCCGGAGAGGATGATGATCTC 1153  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 1043 AGGATGAGAGACTGCTCCCGGAGAGATGATGATGATGATGATGATGATGATGATG 1102  
 QY 1154 CCAATGCTGAGACCGTCAAGGAGCGGAGAGCGGAGAGCTGCAATGGAATCTGAGAGA 1213  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 1103 CGATGCTGAACCGGACAGGAGCGGAGCGGAGAGCGGAGAGCTGCAATGGAATGAGAG 1162  
 QY 1214 TCAACCAATGCGGGGGGTGCTGCTTGTATACCAACCGGAGCGGAGAGGATGAGAG 1273  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 1163 TGAC-----GAGCGCGGCTCCCGGAGCGGAGCGGAGAGCTGCAATGGAATGAGAG 1213  
 QY 1274 AGCACACAGAGATATCCGATGAGTGAAGCTTCCAGAGGATGCTTCCACAGCCCTC 1333  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 1214 AGCACAGGACAGAGATATCCGATGAGTGAAGCTTCCAGAGGATGCTTCCACAGCCCTC 1273

QY 1334 TAAGACAGCCGAGGAGTCCGGTCTTTCTTTTCACCGGAGCCGGGGGCTGAGATGAG 1393  
 Db 1274 TCGACACCCCGG----- 1286  
 QY 1394 CTCGAGCGGGGGCTCCCGACTTCCAAAGCAGACGCTGCTTCTGGGGCCCGAGG 1453  
 Db 1287 ----- 1286  
 QY 1454 GTGGGGGGCGGGGGAGCAGCCCGCCCGAGTCCCGCTCCACACCCCTGCCCGGCC 1513  
 Db 1287 -----GGTACCCCTCACCCCTCACCAAGGGGAGTCCCTCCCC----- 1326  
 QY 1514 CCCAGGCTCCCGCGCTCTCTGGGGAGCCCGCTTCCACTCCGCTCTGACAGCCGCC 1573  
 Db 1327 -----ACCCCGAAGGGAGCAGCTCTCCACAGCCCA 1357  
 QY 1574 GGGCAGCTCCACCGGAGCCCGGGGACACACACACCCCGCGGGGCGCTG 1633  
 Db 1358 AGGAGAGCCCGGGCTGGACAGCCCAACCCCGAGCCCGCTCCAGCCCG-----ACGCTG 1411  
 QY 1634 GGGGAGCCGCTGGAGAGTCTCTCAATCTCCAGCAGAGCTTCTGGGCTCCCTC 1693  
 Db 1412 GAGGGGGCTGGAGGGCGGCTCAATCTCAATCAGAGAGCTTCTGGGCTCACCC 1471  
 QY 1694 GCTTTCACGGCGAGATGAGGTCCTTACCGGTGAGAGATGTCAGCTTGGAGCCAG 1753  
 Db 1472 GCTTTCACCGCGGAGAACTGCAAGTCCGAGCGGAGAGATGTCACCTGAGACAG 1531  
 QY 1754 AGTCTCCCGGAGAGTGGCAAAAGCTCTGTTCTGGGAGCTTCTCTTGGAGAA 1813  
 Db 1532 AGTCTCCCGGAGAGTGGCGAAGAGTCTGTTGGGAACTTCAAGCCGAGAGAG 1591  
 QY 1814 AAGACAAATATCTCTGCTGCTAAGAGCAAACTCTGAGCAGATCAAGCAGATCG 1873  
 Db 1592 AGGAGAGATCTTCTGCTGCTCAAAAGCAAACTCTGAGCTCAATGAGCTGACATCG 1651  
 QY 1874 TCGATGCTTCTGCTGATGCTCCAGCTGAGTCAAGTCTGTCAGACCACTTCA 1933  
 Db 1652 TCGACGCTTCTGCTGATGCTCCAGTCTGAGCAGCAGAGCTATCTCCAAAGAGCTTCC 1711  
 QY 1934 GGGCGAGTACAAAGGAGTGGCGGCGCTCTGCTTCCAAAGCCGCTCGCTTCAG 1993  
 Db 1712 GGGCGAGTACAAAGGAGTGGCGGCGCTCTGCTTCCAAAGCCGCTCGCTTCAG 1771  
 QY 1994 TGGACATCAGCTCTCTGAGAGTCTCAAGCCCTCCCGCAGCGGAGCGGAGAGT 2053  
 Db 1772 TGTATATCAGCTCTGAGAGTCTCAAGCCCTCCCGCAGCGGAGCGGAGAGT 1813  
 QY 2054 GTGGCAGTCTCTGCTGATGCTTCAATCTGAGTCCAGCGGCTGTTCAAGCAG 2113  
 Db 1814 AGGCACTTCTGCTGATGCTTCAAGTCTGAGGCTCCAGCCCGCTGCTTCAAGAGG 1873  
 QY 2114 TGGTGGAGACATCCAGGAGCAGCTCTGAGCAGTCAATGACAGCCCTCCGTCAGG 2173  
 Db 1874 TGGTGGAGACATCCAGGAGCAGCTCTGAGCAGCAGACAGCCGCTGCGGCCAGCACT 1933  
 QY 2174 TGGCAGA 2180  
 Db 1934 TGTGAGA 1940

KW cellular proliferation disorder; cellular differentiation disorder;  
 KW metastatic; haematopoietic disorder; leukaemia; immune disorder;  
 KW inflammatory disorder; arthritis; autoimmune disease; diabetes mellitus;  
 KW psoriasis; Crohn's disease; cardiovascular disease; virus; pain;  
 KW gene therapy; ss.  
 OS Homo sapiens.  
 XX  
 XX  
 Key Location/Qualifiers  
 FT CDS 51..2075  
 FT /tag= a  
 FT /product= "protein kinase 2246"  
 PN  
 XX WO200181588-A2.  
 XX  
 XX 01-NOV-2001.  
 PD  
 XX 25-APR-2001; 2001WO-0513784.  
 PF  
 XX 25-APR-2000; 2000US-199391P.  
 PR  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA  
 XX  
 PI Meyers R;  
 DR WPI; 2002-049281/06.  
 DR P-PSDB; AAM47830.  
 PT  
 PT New protein kinase nucleic acid and polypeptide molecules, designated  
 PT 2246, useful for diagnosing, preventing or treating cancer or a  
 PT cellular proliferation/differentiation disorders, e.g. carcinoma,  
 PT sarcoma or leukaemias  
 PS  
 XX  
 XX  
 Claim 1: Fig 1: 111pp; English.  
 CC The invention relates to the human protein kinase 2246 gene and the  
 CC the isolated encoded polypeptide with cytostatic, immunomodulator,  
 CC anti-inflammatory, analgesic and cardiovascular activity. The 2246  
 CC nucleic acid and polypeptide are useful for diagnosing, preventing or  
 CC treating a subject having cancer or a cellular proliferation and/or  
 CC differentiation disorder or at risk of developing cancer or a cellular  
 CC proliferation and/or differentiation disorder. In particular, the  
 CC disorder includes carcinoma, sarcoma, metastatic or haematopoietic  
 CC disorders (e.g. leukaemia) or cancers of the lung, breast, thyroid, head  
 CC neck, prostate or genito-urinary tract. The 2246 nucleic acid and  
 CC polypeptide are also useful for treating immune disorders, e.g.  
 CC inflammatory (e.g. respiratory inflammation or arthritis), autoimmune  
 CC disease (e.g. diabetes mellitus, psoriasis, Wegener's granulomatosis,  
 CC Crohn's disease or Grave's disease), for treating cardiovascular  
 CC diseases, endothelial cell disorder, viral diseases or pain. The nucleic  
 CC acid and polypeptide are also useful for evaluating the efficacy of a  
 CC treatment of cancer or a cellular proliferation and/or differentiation  
 CC disorder. The nucleic acid is useful for gene therapy. The present  
 CC sequence is that of the 2246 encoding cDNA.  
 SQ  
 Sequence 2217 BP; 476 A; 715 C; 682 G; 344 T; 0 other;  
 Query Match 41.2%; Score 961; DB 24; Length 2217;  
 Best Local Similarity 70.1%; Pred. No. 1.1e-198;  
 Matches 1477; Conservative 0; Mismatches 435; Indels 195; Gaps 6;  
 QY 74 CCCAGCAGCCCAATATGTTGGGCGCTATCGCTGAGAGAGCGTGGCAAGAGCAGA 133  
 Db 79 CGCAGCAGCGCAGATGTTGGGCGCTACCGCTGAGAGAGCGTGGCAAGGGCAGA 138  
 QY 134 CAGGGCTGTTAACTCGGGGCTCAGTCACTACGCGGTGAGAGTCCGATCAAGATCG 193  
 Db 139 CAGGCTCTGTAAGCTGGGGGTTCACTGCTCACTGCGCAAGAGTGGCCATCAAGATCG 198  
 QY 194 TGAACCGGAGAGAGTGTGAGTCTGCTGATGAAGGTGAGAGGAGATCGCCATCC 253  
 Db 199 TCAACCGTGAAGAGTCAACGAGTGTGCTGATGAAGGTGAGAGGAGATCGCCATCC 258

QY	254	TGAACCTCATGAAACACCCACATGTCCTCACTCAACCTCCAGACAGCTCTCGAGAACAAATGAT	313
Db	259	TGAAGCTATTGAGACACCCCCACAGTCTTAAGACTGCACACACCTTTATGAAAACAAAAAT	318
QY	314	ATTGTGACCTGGTTTGTGAGACAGCTCGGGGGGTGAGCTATTGACACTACCTGTAAAGA	373
Db	319	ATTGTACTGGTGTGTAGAAACACAGTGTCAAGTGTGTAGAGTCTTCGACTACCTGTGTAGA	378
QY	374	AGGGAGACTGACGCCCAAGAGGGGCCGAAAGTTCTTCCGCCAGATTGTGTCTGCGCTGG	433
Db	379	AGGGAGGCTGACGCTTAAGGAGGCTCGGAAGTTCCTCCGCGAGACATCTGCGCGCTGG	438
QY	434	ACTTGTGCACAGCTACCTCACTGTGCACAGAGACCTTAAGGCCCGGAACCTGCTTTTGG	493
Db	439	ACTTGTGCACAGCCACCTCATATGTGCACAGGAGATCTGAACCTGAAACCTCTGCTGTGG	498
QY	494	ATGAGAAAAACAACTCCGATTTGGAGAGATTTGGGATGTGCGATGTGCTCTCTGCAGGTGGGGACA	553
Db	499	ACGAGAAAGAACAACTCCGATTCGGAGATTTGGCATTTGGCGTCCCTGCAGAGTTGGGGACA	558
QY	554	GCCCTCTGGAGACCAAGCTGCGGGGTCCCCCATTTATGCTGTCCAGAGTGATTTAAGGGGG	613
Db	559	GCCCTTTGGAGACCAAGCTGTGGGTCCCCCACTACCGCTGCCCCGAGAGTGATTCGGGGGGG	618
QY	614	AAAAATATGATGCGCCCGCGGCGAGACATGTGAGCTGTGAGATCACTCTTGGCCCTGC	673
Db	619	AGAAATATGATGACGCGCGGAGGCGGAGCTGTGAGCTGTGCGCGCTCATCTGTGTGCTCTTGC	678
QY	674	TCTGTGGGGGCTGTGCCCTTTGATGTAGACGAACTCCGCCAGCTGCTGAGAAAGTGAAC	733
Db	679	TGTTGGGGGCTGTGCCCTTCGAGAGATGAACTTTGGGACAGCTGTGTGAGAAAGTGAAC	738
QY	734	GGGGGCTCTTCCACATGCCCCCACTTCATTTCTCCAGATTGCCAGAGCTCTCTGAGGGGAA	793
Db	739	GGGGGCTGTTCACATGCGCGCACTTTATCCCGCCCACTGCCAGAGCTCTCTAGGGGGACA	798
QY	794	TGATGAGATGAGGCCCGGAAAAAGGCTCAAGTGTGAGAAATTCGAAACATCTGTGT	853
Db	799	TGAGGAGATGAGCGCGCGGACCGCGCTTCACGTTAAGACCACTTACGAAACACATATAGT	858
QY	854	ACCTTAGCGCGGAAACACAGACGCCAGACCCGTGCTGTGAGCCAGCGCCGCGCGGCTAG	913
Db	859	ATTATGAGGGGCAAGAAATGAGCCGAAAC-----AAGGCAAGCCATTTCTTCGCAAGGTGC	912
QY	914	CCATGCGGAGCTGCCATTCAAACGAGAGCTGGACCCCGACGTCCTAGAGAGACTGGCAT	973
Db	913	AGATCCGCTCGCTGCCACCTGTGAGAGCATGCAACCCGACGHTGTGAGACATGCACT	972
QY	974	CACCTGGGCTGCTTCAAGGACCCGCGAGAGGCTGCATGCGGAGCTGCCGATGTGAGAGGAGA	1033
Db	973	CACCTGGGCTGCTTCCGAGACCCGCAACAACTGCTGTGAGGCACTGTCTCCGAGAGGAGA	1032
QY	1034	ACCAAGAAAAAGTATATATATATCTGCTTTGTGATGGAAAGAGCGGTATCCAGAGCTGTG	1093
Db	1033	ACCAAGAGAAAGTATATATATCTCTCTCTCTGTGACCCGGAAAGAAAGGTATACCCGAGCCAGG	1092
QY	1094	AGGACCGAGACCTGCCCTCCCGGAATGATGTGACCCCCCGCCGGAAGCGTGTGATTTCTC	1153
Db	1093	AGGATGAGAGACCTGCCCGCCCGGAAAGAGATAGAACCTCCCGGAAAGCGTGTGAGACTGCC	1152
QY	1154	CCATGCTGAGCCGTACAGGGGAAAGCGCGACACAGAGGGAAAGTCCATGGAAGTCTTGACA	1213
Db	1153	CGATGCTGAACCGGACCGCGCAAGCGCGGCGCAGAAACGCAAAATCATATGAGAGTGTCTCAGCG	1212
QY	1214	TCACGAGATCCGCGGGGTGTGGGTCCCTGTACCCACCGCAGCGGCGCTTGGAGATGGGCC	1273
Db	1213	TCAC-----GGAAGCGCGGTCCCGGTGCTGTGCGCGCGGCGCATTTGAGATGGGCC	1263
QY	1274	ACGACAGCCAGAGATCCCTAGCGCTAGTGAAGCTTCCACAGGGGTCTGTCTCTCAGCCCTC	1333
Db	1264	ACGACAGCCAGAGGTCTGTGGTTCATAGAGGAGTCTCTCTCAGGCGCTTTTCACACAGCCGCAC	1323
QY	1334	TAAAGCAAGCCAAAGAGTCCGGTCTTTTCTTTTCACCGGAGCCGCGGGGCTGTGAGATGAG	1393

Db	1324	TCACGACGCCCCG	-----	1336
QY	1394	CTCAGAGCGGGGGCTCCCCGACTTCCAAAAAGCAGACGTCGCTTCTGGGGCCCCAGGG		1453
Db	1337	-----	-----	1336
QY	1454	GTGGGGGCGCGGGGAGCAGCCCCCGCCCTCAGTGGCCGCTCCACACCCCTGCGCGGC		1513
Db	1337	-----	-----	1376
QY	1514	CCCCAGGCTCCCCGCGTCTCTGGCGGGAGACCCCTTGCACTGCGCCCTGTCACACAGCCCC		1573
Db	1377	-----	-----	1407
QY	1574	GGCCAGTCCACCCGGGAGACCCCGGGGCAACACACCCACCCAGCCCCCGCGGTGGCGTCG		1633
Db	1408	AGGAGAGCGCGGCTGGGACGCGCCAAACCCAGCGCCCGCTCGACGCCC	-----	1461
QY	1634	GGGAGCGCGCTGGAGGAGTGTCTCAACTCATCCGCAACAGCTTCTTGAGCTCCCTC		1693
Db	1462	GAGGGGTGCTCGAGGGCGCGGGCTCAACTCAATCAAGAACAGTTTCTGGGCTCACCCC		1521
QY	1694	GCTTCAACCGCGGAGATGTAGGTGCTCCCTACCGGTGAGGATGTCCAGTTGACGCGAG		1753
Db	1522	GCTTCCACCGCGGAAACTGCAACTTCCAGCGCCGGAGGAGATGTCCAACTGACACAG		1581
QY	1754	AGTCTCTCCCGAGCTGGCAAAACGCTCTGGTTCGGGAACTTCATCTCTTGGACAAAG		1813
Db	1582	AGTGTGCTCCGAGACTGGCGAAGAAAGTCTGGTTGGGAACCTTATGAGCTGGAGAAAG		1641
QY	1814	AAGAACAAATTTCTCTGTGCTTAAGGACAACTCTCAGCAGCATTAAGACAGCATCG		1873
Db	1642	AGGAGCAGATCTTGTGTATCATCAAAAGCAAACTCTGAGCTCCATCAAGGCTGACATCG		1701
QY	1874	TCCATGCTTCTGTGATGCCCGCGGTGACAGTGTGCTGTACAGACCAAGCTTCA		1933
Db	1702	TGCAGCGCTTCTGTGATTTCCAGTCTCACGCCACAGGTCATTTCCAAACAGAGCTTCC		1761
QY	1934	GGGCGGATACAAAGCCAGTGGCGGCCCTTCCCTTCCAAAGCCCGTCCGCTTCCAGG		1993
Db	1762	GGGCGGATACAAAGCCAGGGGGGCGCAGCGGTTCAGAAACCGGATCAAGTTCAGG		1821
QY	1994	TGACATCAGTCTCTCTGAGGGGTCCAGAGCCCTCCCGCGCAGGGAGCGCAGCGGAGTG		2053
Db	1822	TTGATATCACCCTACACGAGGGGT-----	GGGGAGGCCACGAAGAGAGA	1863
QY	2054	GTGCGATTAATCCGTACCTTCACTCATCTGAGGGGTCCAGCGCGTGGTTCAAGGAG		2113
Db	1864	ACGGCATCTACTCGCTCACTTCAACCTGCTCTCAGGCGCCAGCGCGTCTCAAGAGGG		1923
QY	2114	TGGTGGAGACCATCCAGGAGCAGCTCCGAGACATCAATGACACAGCCCTCCGTGAGGCCC		2173
Db	1924	TGGTGGAGACCATCCAGGAGCAGCTGCTGAGACACACAGACCGCGCTGGCGCCAGCACT		1983
QY	2174	TGGCAGA	2180	
Db	1984	TGTCAGA	1990	
RESULT 5				
AAD34315				
AAD34315				
AAD34315				
AAD34315				
16-JUL-2002				
Human PKIN-18 cDNA.				
Human: kinase, enzyme; PKIN-18 protein; immune system disorder; anaemia;				
acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;				
asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;				

KW AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;  
 KM leukemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;  
 KM Down's syndrome; gene therapy; protein therapy; cytosolic; gene; ss.  
 XX Homo sapiens.

XX Key Location/Qualifiers  
 FT CDS 1.1995  
 FT /tag= a  
 FT /product= "Human PKIN-18 protein"

MO200218557-A2.

07-MAR-2002.

31-AUG-2001; 2001WO-US27219.

PR 31-AUG-2000; 2000US-229873P.  
 PR 08-SEP-2000; 2000US-231357P.  
 PR 14-SEP-2000; 2000US-232654P.  
 PR 22-SEP-2000; 2000US-234902P.  
 PR 29-SEP-2000; 2000US-236499P.  
 PR 06-OCT-2000; 2000US-238389P.  
 PR 13-OCT-2000; 2000US-240542P.

XX (INCY-) INCYTE GENOMICS INC.

XX Bandman O, Nguyen DB, Walla NK, Hafalia AJA, Yao MG, Gandhi AR;  
 PI Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;  
 PI Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YT;  
 PI Azimtal Y, Burdill JD, Marcus GA, Zingler KA, Lu DM, Lal PG;  
 PI Rankumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;  
 PI Burford N;

XX WPI; 2002-329769/36.  
 DR P-PSDB; AAE31723.

XX New human kinases, useful for diagnosing, treating or preventing immune  
 PT system disorders (e.g. Crohn's disease), neurological disorders (e.g.  
 PT epilepsy), or cell proliferative disorders (e.g. cancers such as  
 PT leukemia or lymphoma)

XX Claim 97; Page 212-213; 218pp; English.

XX The present invention relates to human kinases (PKIN) and polynucleotides  
 CC encoding such proteins. PKIN sequences of the invention are useful for  
 CC diagnosing, treating or preventing disorders associated with aberrant  
 CC expression of PKIN, particularly immune system disorders (e.g. acquired  
 CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,  
 CC anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-  
 CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers  
 CC such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma),  
 CC and developmental disorders (e.g. Down's syndrome). They are also used  
 CC in gene therapy and protein therapy. The present sequence is a cDNA  
 CC encoding human PKIN-18 protein.

XX Sequence 2647 BP; 525 A; 885 C; 789 G; 448 T; 0 other;

XX Query Match 41.0%; Score 955.8; DB 24; Length 2647;  
 XX Best Local Similarity 70.0%; Pred. No. 1.6e-197;  
 XX Matches 1473; Conservative 0; Mismatches 437; Indels 195; Gaps 6;

QY 88 TATGTGGGCCCCATCGGCTGGAGAGAGCGTGGGCAAGAGACAGACAGGCGTGTAAA 147  
 DB 31 TATGTGGGCCCCATCGGCTGGAGAGAGCGTGGGCAAGAGAGCGTGTGTGAAG 90  
 QY 148 CTGCGGGTCCAGTGCATGAGGAGGTCAGAGGTCGCGCATCAAGATGTAACCGGAGAG 207  
 DB 91 CTGCGGGTCCAGTGCATGAGGAGGTCAGAGGTCGCGCATCAAGATGTAACCGGAGAG 150  
 QY 208 CTGCGGAGTGCATGAGGAGGTCAGAGGTCGCGCATCAAGATGTAACCGGAGAG 267  
 DB 151 CTGAGGAGTGCATGAGGAGGTCAGAGGTCGCGCATCAAGATGTAACCGGAGAG 210

QY 268 CACCACATGCTCTCAAGCTCCAGAGCTCTACGAGAACAGAAATATTGTACCTGGTT 327  
 DB 211 CACCCACAGCTCTAAAGCTCAGAGCTTATGAAAACAAAAAATATTGTACCTGGTT 270  
 QY 328 CTGAGAGAGCTGGGGGGGAGAGCTATTGACATCTCTGTGTAAAGAGGAGACTGAC 387  
 DB 271 CTGAGAGAGCTGGGGGGGAGAGCTATTGACATCTCTGTGTAAAGAGGAGACTGAC 330  
 QY 388 CCGAAGAGAGCCGAAAGTCTTCCGAGATGTGTGTGGCTGGAGCTTCCACAGC 447  
 DB 331 CTAAGAGAGGCTGGAGAGTCTTCCGAGATGTGTGTGGCTGGAGCTTCCACAGC 390  
 QY 448 TACTCATCTGCCACAGAGACTTAAAGCCGAGAACTGCTTTTGGATGAGAAAAAC 507  
 DB 391 CACTCATATGCTCCACAGAGATCTGAAACCTGAAACCTCTGTGTGACAGAGAAAC 450  
 QY 508 ATCCGATTCAGAGCTTCGAGATGGCTCTCTGACAGGTGGGGAGACGCTCTGAGACC 567  
 DB 451 ATCCGATTCAGAGCTTCGAGATGGCTCTCTGACAGGTGGGGAGACGCTCTGAGACC 510  
 QY 568 AGCTGCGGTCGCCCATTTATGCTGTCCAGAGGTGATTAAGGGGAGAAATATGATGGC 627  
 DB 511 AGCTGCGGTCGCCCATTTATGCTGTCCAGAGGTGATTAAGGGGAGAAATATGATGGC 570  
 QY 628 CGCGGGGACAGATGTGAGAGCTGTGAGTCACTCTTCCCTGCTGCTGGGGGCTCTG 687  
 DB 571 CGGAGGCGGAGCTGTGAGAGCTGTGAGTCACTCTTCCCTGCTGCTGGGGGCTCTG 630  
 QY 688 CCTTTGATATGACAGCAACCTCCGACAGCTGTGTGAGAGAGTGAAGGGGCTTCCAC 747  
 DB 631 CCTTTGATATGACAGCAACCTCCGACAGCTGTGTGAGAGAGTGAAGGGGCTTCCAC 690  
 QY 748 ATGCCCATTTCAATCTCCAGATTCAGATTCGAGAGCTCTCTGAGGGAGATGAGAGT 807  
 DB 691 ATGCCCATTTCAATCTCCAGATTCGAGAGCTCTCTGAGGGAGATGAGAGT 750  
 QY 808 CCGGAAAAAGGCTCACTGTGAGAGCAAAATTCAGAAATCTTGTATCAGGGGGGAA 867  
 DB 751 GCGGACGCGGCTCAGAGTACAGACATTCAGAAACATATGTATATAGGGGGCAAG 810  
 QY 868 CAGGAGCAAGCCGCTCTGAGAGCCAGCCCTGGCGCGGGATAGCATGCGAGCCTG 927  
 DB 811 AATGAGCCGAGAAC-----AGAGCAGCCATTCCTCTGCAAGGTGACAGATCGCTG 864  
 QY 928 CCATTCAGAGAGAGCTGAGACCCGAGCTCTTACAGAGATGATCATCTGGGCTGTC 987  
 DB 865 CCAGGCTGAGAGAGATGAGACCCGAGCTCTTACAGAGATGATCATCTGGGCTGTC 924  
 QY 988 AGGAGCCGAGAGAGCTGATCGGAGCTGCGAGTGTGAGAGAGAGAGAAAGATG 1047  
 DB 925 CGAGACCGCAAGAGCTGTGAGAGCTGCTCTGCGAGGAGAGAGAACAGAGAAATG 984  
 QY 1048 ATATATATGCTTTGATGAGAGAGAGAGAGAGAGATTCAGAGCTGTGAGAGAGAG 1107  
 DB 985 ATATATGCTTTGATGAGAGAGAGAGAGAGAGATTCAGAGCTGTGAGAGAGAG 1044  
 QY 1108 CCTCCCGAATGATGTTGAGACCCCGGAGAGGCTGTGATTCATCTGAGAGCCT 1167  
 DB 1045 CCCCCCGAGAGAGATGAGACCTCCCGGAGAGGCTGTGATTCATCTGAGAGCCT 1104  
 QY 1168 CACGGAGAGCGGAG 1227  
 DB 1105 CACGGAGAGCGGAG 1155  
 QY 1228 GGTGTGTGCTCCCTGTACACACCGAGGCTGTGAGAGATGAGAGAGAGAGAG 1287  
 DB 1156 GACGGGCTCCCGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1215  
 QY 1288 TCCCGTACGCTAGTGTGAGAGCTCAGAGGCTGTGAGAGAGAGAGAGAGAGAGAG 1347  
 DB 1216 TGTGGTTCATGAG 1270

QY 1348 AGCCGCTCTTTCTTTTCCACCGGACCGGGGCTGAGATGAGGCTCAGCGGGGGC 1407  
 DB 1271 ----- 1270  
 QY 1408 TCCCGACTTCAAAAGAGAGAGCTGCTTCTTCGGGGCCCCAGGGGGTGGGGGGCCGGG 1467  
 DB 1271 ----- 1277  
 QY 1468 GAGCAGCCCCCGCCCGAGTGGCCGCTGCAGACCCCTGCCGGGCCCCAGGCTCCCG 1527  
 DB 1278 GACCCCTCAGCCCTCAGCAAGGGGAGTCCCGTCCG----- 1314  
 QY 1528 CGCTCTCTGGGGGAGACCCCTTGCACTCGCTCTGCACACGCCCGCCGAGTCCAC 1587  
 DB 1315 ----- 1359  
 QY 1588 GGGACCCCGGGGACACACACCCCGGCGGCGGTGGGCGGGAGCGCGCTGG 1647  
 DB 1360 GGCAGCCCAACCCCAAGCCCGGCTCCAGCCG-----AGCTCGGAGGGGTGCTGG 1413  
 QY 1648 AGGAGTCTCTCACTCACTCCGACAGGCTTCTGGGCTCCCTGCTTCCACCGCGC 1707  
 DB 1414 AGGGCGGCTCACTCACTCAAGAACAGCTTCTGGGCTCAGCCCGCTCCACCGCGG 1473  
 QY 1708 AGATGCAAGTCTCTCACTCACTCCGAGAGATGTCAGCTTGCAGCCAGAGTCTCCCGGAG 1767  
 DB 1474 AACTGCAAGTCTCGAGCGCGGAGATGTCACCTGACACAGATCTCTCCAGAG 1533  
 QY 1768 CTGGCAAAAGCTCTGCTGGTGGGAACTTCACTCTTGGCAAAAGAACAAATATTC 1827  
 DB 1534 CTGGCAAAAGCTCTGCTGGTGGGAACTTCACTCTTGGCAAAAGAACAAATATTC 1593  
 QY 1828 CTGCTCTAAAGAGCAAACTCTCAGAGCATCAAGCATCAAGCATGCTCTTCTG 1887  
 DB 1594 GTGGTCAATCAAAAGCAAACTCTGAGCTCAATCAAGCTGATGCTGACGCTTCTG 1653  
 QY 1888 TCGATCCCAAGCTCTGAGTCACTGCTGTCAGACAGACAGTTCAGGCGGAGTCAAG 1947  
 DB 1654 TCGATCCCAAGCTCTGAGTCACTGCTGTCAGACAGTTCATCTCCAAAGACTTCCGGCGAGTCAAG 1713  
 QY 1948 GCGAGTGGGGGCGGCTCCGCTTCCAAAGAGCGGCTCCGCTCCAGAGTGGATCAAGTCC 2007  
 DB 1714 GCGAGTGGGGGCGGCTCCGCTTCCAAAGAGCGGCTCCGCTCCAGAGTGGATCAAGTCC 1773  
 QY 2008 TCTGAGGGTCCAGAGGCTCCCGCGAGCGGAGCGGAGGAGTGGATCAAGTCC 2067  
 DB 1774 AGGAGGGT-----GGGAGGCGCAGAGAGAGAGATGATCTACTCC 1815  
 QY 2068 GTCACTTCACTCTCATCTCTGGGTCCAGCGGCTGGTTCAGAGAGTGGTGAACCATC 2127  
 DB 1816 GTCACTTCACTCTCATCTCTGGGTCCAGCGGCTGGTTCAGAGAGTGGTGAACCATC 1875  
 QY 2128 CAGGCAAGCTCTGAGCATCTAGACAGCCCTCCGCTGAGGCGCTGGAGAGAGAG 2187  
 DB 1876 CAGGCGAGTGTGTAGACACACAGACAGCCCGCTCCGCGCAGACTTGTACAGACCACT 1935  
 QY 2188 AAGCG 2192  
 DB 1936 AACTG 1940  
 RESULT 6  
 ABA08296  
 ID ABA08296 standard; cDNA; 906 BP.  
 AC ABA08296;  
 XX  
 XX 11-JAN-2002 (first entry)  
 DE Human HtPOB-1 homologue-encoding cDNA, SEQ ID NO:72.  
 XX Human; cytokine; cell proliferation; cell differentiation; growth factor;  
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;

KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 KW chronic inflammatory condition; proliferative retinopathy;  
 KW atherosclerosis; coronary heart disease; arterial ischaemia;  
 KW bone disorder; osteoporosis; vascular growth disorder;  
 KW tissue regeneration; wound healing; infection; immune disorder;  
 KW cell culture; drug screening; gene therapy; antiinflammatory;  
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
 KW cytosolic; osteopathic; vasotropic; cardiatic; virucide; antibacterial;  
 KW antifungal; vulnery; antilucer; ss.  
 OS Homo sapiens.  
 PN WO200157188-A2.  
 PD 09-ADG-2001.  
 PF 05-FEB-2001; 2001WO-US03800.  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PA (HSE-) HSEQ INC.  
 PI Tang YT, Liu C, Drmanac RT;  
 DR WPI: 2001-457740/49.  
 PS P-PSDB; ABB11052.  
 PT Human proteins and DNA encoding sequences useful for preventing,  
 PT treating or ameliorating a medical condition in a mammalian subject  
 PT e.g. arthritis and cancer.  
 Claim 1: Page 364-365; 1963pp; English.  
 XX  
 XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
 CC invention also relates to vectors and recombinant host cells comprising a  
 CC nucleotide of the invention, methods of producing the novel polypeptides,  
 CC antibodies against the polypeptides, methods of detecting the nucleotides  
 CC or polypeptides in a sample, and methods of identifying compounds which  
 CC bind to polypeptides of the invention. Although novel, many of the  
 CC polypeptides of the invention have homology to known proteins, thereby  
 CC giving an insight into their probable biological activities, and hence  
 CC potential therapeutic applications. The polypeptides of the invention may  
 CC have various activities, including cytokine, cell proliferation or cell  
 CC differentiation activities; stem cell growth factor activity;  
 CC haematopoiesis regulatory activity; tissue growth activity;  
 CC immunomodulatory activity; activin- or inhibin-related activities;  
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
 CC thrombolytic activities; receptor or ligand activities; or may be  
 CC involved in oncogenesis, cancer cell proliferation or metastasis.  
 CC Depending on their biological activities, polypeptides and nucleotides of  
 CC the invention are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g., by protein or gene therapy. Such conditions include  
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
 CC vascular growth. Polypeptides involved with tissue regeneration and  
 CC repair (or nucleic acids encoding them) may be used to promote wound  
 CC healing (e.g., of burns, incisions and ulcers), while those with  
 CC immunomodulatory activities may be used in the treatment of viral,  
 CC bacterial and fungal infections in addition to immune disorders.  
 CC Polypeptides with growth factor activity may be used in cell cultures to  
 CC promote cell growth. For example, such polypeptides may be used to  
 CC manipulate stem cells in culture to give rise to neuroepithelial cells  
 CC that can be used to augment or replace cells damaged by illness,  
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
 CC may also be used in the diagnosis of the above conditions, and in drug  
 CC screening techniques. The present sequence represents a cDNA encoding a  
 CC novel human polypeptide of the invention.

XX Sequence 906 BP; 205 A; 253 C; 267 G; 181 T; 0 other;  
 SQ Best Local Similarity 28.2%; Score 658.8; DB 22; Length 906;  
 Query Match 87.9%; Pred. No. 2.9e-133;  
 Matches 786; Conservative 0; Mismatches 2; Indels 106; Gaps 2;

```

  79 CACGCCCAATATGTGGGCGCCCTATCGGCTGGAGAAGACGCTGGGCGAAGAGACAGACAGG 138
  1 CACGCCCAATATGTGGGCGCCCTATCGGCTGGAGAAGACGCTGGGCGAAGAGACAGACAGG 60
  139 CCGGTTAAATCGGGGCGCCCTATCGGCTGGAGAAGACGCTGGGCGAAGAGATCGTGAAC 198
  61 CCGGTTAAATCGGGGCGCCCTATCGGCTGGAGAAGACGCTGGGCGAAGAGATCGTGAAC 120
  199 CCGGAGAAGCTGTGCGAGTGTGCTGATGAGAGGTGGAGGAGATCGCATCT-GAA 257
  121 CCGGAGAAGCTGTGCGAGTGTGCTGATGAGAGGTGGAGGAGATCGCATCTCGAG 180
  258 GCTCATCGAACCACCCACATGTCTCAAGCTCCACGACGCTTACAGAGAAGAAATATTT 317
  181 GCTCATCGAACCACCCACATGTCTCAAGCTCCACGACGCTTACAGAGAAGAAATATTT 240
  318 ----- 317
  241 TCCCCCGGATGAACATGACATCAGGTCGTCGANTGCTGGCGAGGTTTCACACAGGGAA 300
  318 -----GTACCTGTTCTGGA 332
  301 GCTGTGTCGCCCGGCGCTCATGGAGACCTGCTCAGGCTTCCAGAGTACCTGTTCTGGA 360
  333 GCACGTCCTGGGGGGGAGTGTGATTCGATCAGTCTGTAAGAGAGGAGACATGACGCCAA 392
  361 GCACGTCCTGGGGGGGAGTGTGATTCGATCAGTCTGTAAGAGAGGAGACATGACGCCAA 420
  393 GCGAGGCGCGAAGTCTGTCGCGAGATGTCGCGCTGACCTTGGCCACAGTACTC 452
  421 GCGAGGCGCGAAGTCTGTCGCGAGATGTCGCGCTGACCTTGGCCACAGTACTC 480
  453 CATCTGCCACAGAGACCTAAAGCCCGAAGACCTGTTGGATGAGAAAACAACATCCG 512
  481 CATCTGCCACAGAGACCTAAAGCCCGAAGACCTGTTGGATGAGAAAACAACATCCG 540
  513 CATCTGCCACAGAGTGGCATGGCGCTCCCTGCAGAGTGGGGACACCTCTCGAGACCACTG 572
  541 CATCTGCCACAGAGTGGCATGGCGCTCCCTGCAGAGTGGGGACACCTCTCGAGACCACTG 600
  573 CCGGTCGCCCGCATTTAGCGTGCAGAGGTGATTAGGGGGGAAAATATGATGGCGCGG 632
  601 CCGGTCGCCCGCATTTAGCGTGCAGAGGTGATTAGGGGGGAAAATATGATGGCGCGG 660
  633 GCGAGACATGTGGAGTGTGAGATCATCTTCGCGCTGCTGGGGGCTGCGCCCT 692
  661 GCGAGACATGTGGAGTGTGAGATCATCTTCGCGCTGCTGGGGGCTGCGCCCT 720
  693 TGTATGACGACACCTCCGCCAGCTGCTGAGAGAGTGAACGCGGCGCTTTCACATGCC 752
  721 TGTATGACGACACCTCCGCCAGCTGCTGAGAGAGTGAACGCGGCGCTTTCACATGCC 780
  753 CCACTTCATTCCTCCAGATTCGACAGACCTCCGAGGGGATGATGAGTGGAGCCCGA 812
  781 CCACTTCATTCCTCCAGATTCGACAGACCTCCGAGGGGATGATGAGTGGAGCCCGA 840
  813 AAAAAGCTCAGTCTGAGAGCAATTCGAAAACATCTTGTAGCTAGCGGGAA 866
  841 AAAAAGCTCAGTCTGAGAGCAATTCGAAAACATCTTGTAGCTAGCGGGAA 894
  
```

XX 11-JAN-2002 (first entry)  
 DE Human Ser/Thr kinase homologue-encoding cDNA, SEQ ID NO:700.  
 XX Human; cytokine; cell proliferation; cell differentiation; growth factor;  
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 KW chronic inflammatory condition; proliferative retinopathy;  
 KW atherosclerosis; coronary heart disease; arterial ischaemia;  
 KW bone disorder; osteoporosis; vascular growth disorder;  
 KW tissue regeneration; wound healing; infection; immune disorder;  
 KW cell culture; drug screening; gene therapy; antiinflammatory;  
 KW antisthmatic; antiarthritis; haemostatic; antiarteriosclerotic;  
 KW cytotatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
 KW antifungal; vulnery; antitumor; ss.

OS Homo sapiens.  
 PN WO200157188-A2.  
 XX 09-AUG-2001.  
 PD 05-FEB-2001; 2001WO-US03800.  
 PF 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 XX (HSE-) HSEQ INC.  
 PA Tang YT, Liu C, Drmanac RT;  
 PI WPI: 2001-457740/49.  
 XX P-PSDB: ABB11680.  
 DR Human proteins and DNA encoding sequences useful for preventing,  
 PT treating or ameliorating a medical condition in a mammalian subject  
 PT e.g. arthritis and cancer -  
 XX Claim 1; Page 648; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
 CC invention also relates to vectors and recombinant host cells comprising a  
 CC nucleotide of the invention. Methods of producing the novel polypeptides,  
 CC antibodies against the polypeptides, methods of detecting the nucleotides,  
 CC or polypeptides in a sample, and methods of identifying compounds which  
 CC bind to polypeptides of the invention. Although novel, many of the  
 CC polypeptides of the invention have homology to known proteins, thereby  
 CC giving an insight into their probable biological activities, and hence  
 CC potential therapeutic applications. The polypeptides of the invention may  
 CC have various activities, including cytokine, cell proliferation or cell  
 CC differentiation activities; stem cell growth factor activity;  
 CC haematopoiesis regulatory activity; tissue growth activity;  
 CC immunomodulatory activity; activin- or inhibin-related activities;  
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
 CC thrombolytic activities; receptor or ligand activities; or may be  
 CC involved in oncogenesis, cancer cell proliferation or metastasis.  
 CC Depending on their biological activities, polypeptides and nucleotides of  
 CC the invention are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g., by protein or gene therapy. Such conditions include  
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
 CC vascular growth. Polypeptides involved with tissue regeneration and  
 CC repair (or nucleic acids encoding them) may be used to promote wound  
 CC healing (e.g., of burns, incisions and ulcers), while those with  
 CC immunomodulatory activities may be used in the treatment of viral,  
 CC bacterial and fungal infections in addition to immune disorders.  
 CC Polypeptides with growth factor activity may be used in cell cultures to

CC promote cell growth. For example, such polypeptides may be used to  
 CC manipulate stem cells in culture to give rise to neuroepithelial cells  
 CC that can be used to augment or replace cells damaged by illness,  
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
 CC may also be used in the diagnosis of the above conditions, and in drug  
 CC screening techniques. The present sequence represents a cDNA encoding a  
 CC novel human polypeptide of the invention.

XX Sequence 614 BP; 128 A; 226 C; 161 G; 99 T; 0 other;

Query Match 26.3%; Score 614; DB 22; Length 614;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-123;  
 Matches 614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1718 TCCCTACCGCTGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 1777  
 DB 1 TCCCTACCGCTGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 60  
 QY 1778 GCTCCGCTGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 1837  
 DB 61 GCTCCGCTGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 120  
 QY 1838 AGGACAACTCTCAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 1897  
 DB 121 AGGACAACTCTCAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 180  
 QY 1898 GCTCCGCTGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 1957  
 DB 181 GCTCCGCTGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 240  
 QY 1958 GCTCCGCTGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 2017  
 DB 241 GCTCCGCTGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 300  
 QY 2018 CAGAGCTCTCCCGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 2077  
 DB 301 CAGAGCTCTCCCGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 360  
 QY 2078 CAGAGCTCTCCCGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 2137  
 DB 361 CAGAGCTCTCCCGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 420  
 QY 2138 TCCCTACCGCTGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 2197  
 DB 421 TCCCTACCGCTGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 480  
 QY 2198 AGGACAACTCTCAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 2257  
 DB 481 AGGACAACTCTCAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 540  
 QY 2258 CAGAGCTCTCCCGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 2317  
 DB 541 CAGAGCTCTCCCGAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 600  
 QY 2318 AGGACAACTCTCAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 2381  
 DB 601 AGGACAACTCTCAGAGATGTCAGCTTGAACGCGAGAGTCTCCCGAGAGTGGCAAAAC 660

RESULT 8  
 ABK70216  
 ID ABK70216 standard; cDNA; 512 BP.

XX ABK70216;  
 XX 15-JUL-2002 (first entry)

DE Human lung cancer associated cDNA SEQ ID 87.

KW Human; ss; gene; lung cancer; cytostatic; tumour; vaccine.

OS Homo sapiens.  
 XX

PN WO200224057-A2.  
 XX 28-MAR-2002.  
 PD 20-SEP-2001; 2001WO-US42232.  
 PP 22-SEP-2000; 2000US-234837P.  
 PR 10-OCT-2000; 2000US-239440P.  
 PR 29-JUN-2001; 2001US-301928P.  
 PA (CORI-) CORIXA CORP.  
 XX Benson DR, Mohamath R, Lodes MJ;  
 PI WPI; 2002-372001/40.  
 DR New tumour lung proteins and nucleic acids encoding the proteins, useful  
 PT as vaccines and for treating, preventing, diagnosing or monitoring lung  
 PT cancer.  
 PS Claim 1; Page 132-133; 189pp; English.

The invention relates to an isolated polynucleotide comprising a sequence  
 selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),  
 or their fragments, homologues, variants or complements and their encoded  
 polypeptides. Also included are an expression vector comprising the  
 polynucleotide operably linked to an expression control sequence; a host  
 cell transformed or transfected with an expression vector of; an isolated  
 antibody, or its antigen-binding fragment that specifically binds to the  
 polypeptide; a method for detecting the presence of a cancer in a  
 patient; a fusion protein comprising at least the polypeptide; an  
 oligonucleotide that hybridises to the polynucleotide under moderately  
 stringent conditions; a method for stimulating and/or expanding T cells  
 specific for a tumour protein; an isolated T cell population comprising T  
 cells prepared from the method of above; a composition comprising a first  
 component consisting of carriers and immunostimulants, and a second  
 component selected from the polynucleotides, proteins, antibodies, fusion  
 proteins, T cell populations and antigen presenting cells expressing the  
 polypeptide; methods for stimulating an immune response or treating  
 cancer in a patient by administering the composition and diagnostic kits  
 comprising at least one of the oligonucleotide of, or an antibody and a  
 detection reagent consisting of a reporter group. The polypeptides and  
 polynucleotides are useful as vaccines for the treatment or prevention of  
 lung cancer, and for diagnosis and monitoring of such cancer. The  
 polynucleotide, polypeptide and antigen presenting cells can be  
 used to stimulate or expand T cells specific for a tumourous protein.  
 CC The polynucleotides may be used as probes or primers for nucleic acid  
 CC hybridisation, and in the preparation of ribozyme molecules for  
 CC inhibiting expression of tumour polypeptides and proteins in tumour  
 CC cells. The present sequence is one of the 183 lung cancer associated  
 CC polynucleotides.

XX Sequence 512 BP; 108 A; 149 C; 166 G; 89 T; 0 other;

Query Match 21.9%; Score 512; DB 24; Length 512;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-101;  
 Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 519 AGACTTCGCGATGCGGCTGCGAGTGGGGGAGACAGCTCTCGAGACAGCTGGGGTC 578  
 DB 1 AGACTTCGCGATGCGGCTGCGAGTGGGGGAGACAGCTCTCGAGACAGCTGGGGTC 60  
 QY 579 CCCCATTATGCTGTCAGAGTGTAAAGGGGAAATATGATGAGCCCGGGGAGA 638  
 DB 61 CCCCATTATGCTGTCAGAGTGTAAAGGGGAAATATGATGAGCCCGGGGAGA 120  
 QY 639 CAGTGTGAGTGTGAGTATCTCTTGGCCCTGCTGTGGGGGCTTGCCTTTGATGA 698  
 DB 121 CAGTGTGAGTGTGAGTATCTCTTGGCCCTGCTGTGGGGGCTTGCCTTTGATGA 180  
 QY 699 CGACAACCTCCGCGAGTGTGAGAGTGAAGAGGGGGGCTCTTCACATGCCCCACTT 758  
 DB 181 CGACAACCTCCGCGAGTGTGAGAGTGAAGAGGGGGGCTCTTCACATGCCCCACTT 240









CC interaction. They can also be used to identify agents useful in the  
CC prevention, treatment and diagnosis of neurological diseases, including  
CC stroke, multiple sclerosis, Parkinson's disease, amyotrophic lateral  
CC sclerosis, tumors, especially cardioma, immunological disorders,  
CC including autoimmune diseases, atopy, viral (including human immune  
CC deficiency virus) infections, acute or chronic leukocytic or myeloid  
CC leukemias, primary chronic polyarthritis, Crohn's disease and ulcerative  
CC colitis. The present sequence is the human 955\_b coding sequence.

Sequence 3392 BP; 726 A; 1099 C; 972 G; 595 T; 0 other;

Query Match	Score	DB	Length
Best Local String	11.38;	264.2;	DB 24;
Best Local String	57.69;	130.47;	

Matches 473; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

QY	32	GTCTCCCGGCTTACACACTCCCGCACCGCCCGCACCGCCAGACCCCAATATG	91
Db	170	GCTACTG6GTGCCCGTTGGCCGGAATCTCATGCGCTCTGTCCGAGAGACGCCACG	223
QY	92	TGGGCCCTTATCGGCTGTGAGAGACGCTGGGCAAGACACAGAGGCTGGTTAACTCG	151
Db	230	TGGGCAACTACCGGCTCGGTGTGAGACCAATTGGGAAGGGCACTTTGCCAANTCAAGCTGG	283
QY	152	GGGTCCACTGCATACAGGGGTGAGAGGTCCGCTCAAGANTGTGAACCGGGAGAAGCTGT	211
Db	290	CTCGGCAACATCTCACTACGTGTGGGAGAGGTGGCCATCAAGATTATGACAAACCCAGCTGA	343
QY	212	CGAGTGGGTCTGATGAAGGTGGAGGGGAGATCGCATCTGAAGTCATGCAACCC	273
Db	350	ATCCAGACAGGCTCTCAGAAAGCTGTGTCCGAGAGATGCCATATATGAAGGCTTAACACC	403
QY	272	CACATGTCTTCACAGCTCCACAGAGCTACAGAAACAGAAATATTGTACTGGTTCTGG	333
Db	410	CCAACATCTGTAACCTCTTTGAGTGTGACTGAGAGAGACGCTGTACTGGTATGG	463
QY	332	AGCAGCTCTCGGGGGGGTGAAGTATTGCATCTCTGGTGAAGAGAGGAGACTACGCCCA	393
Db	470	AGTACGGAAGTGTGGAGAAAGTGTTAACATCTCTGTGTCAGTGGCCGACATGAAGAGAGA	523
QY	392	AGGAGGCGGAAAGTCTTCGCGGAGATGTGTGGCGTGGCACTTGGCAGAGCTACT	453
Db	530	AGGAAAGCTCAGAGCCAAAGTCTCGACAGATTTGTTGGCTGTGCATATTGTTCACAGAAA	583
QY	452	CCATCTGCCACAGAGACCTTAAGCCGAGAAACCTGTTTGGATGAGAAAAACAATCC	513
Db	590	ATATTGTACACAGGAGCCTGAAAGGCTGTGAAGACCTCTTGCTGGATGGCCGAGCCAAATCA	643
QY	512	GCATTGCAAGACTTGGGCAATGGCGTCCCTGCAGGTGGGGGACAGGCTCCTGGAACCAAGT	573
Db	650	AGATTGTGACTTGTGGCTTCACAGCAAGATTCAAGCTGGGATCGAACTGTGACACGTTCT	703
QY	572	GGCGGTCCTCCCATTTATGCGTGTCCAGAGGTGATTAAGGGGAGAAATATGATGGCCGC	633
Db	710	GGCGGAGACCCCCCAATTATGCCGCCGCCGAGAGCTGTTTCAGGGCCAGAAATAGACGGGCGG	763
QY	632	GGGCAGACATGTGAGAGCTGTGAGATCACTCTTCGCGCTGCTGTGGGGACTGTGCGCT	693
Db	770	AGGTGGACATCTGGAGGCTGGGAGTCACTCTGTACACCTCTGTACAGCGCTCCCTGCGCT	823
QY	692	TGTATGACGACAACTCCGCGAGCTGTGTGGAGAAAGTGAACGGGGGCGTCTTCCACATGC	753
Db	830	TGCAAGGGCACAACTCAAGAGAGTGGGGAGCGAGTACTCAGAGGGAAGTACCGGGGTCC	883
QY	752	CCGACCTTCATCTCCGAGATTGGCCAGAGCCCTCGTAGGGGAATATGATGAATGGAGCCCG	813
Db	890	CTTTCTTCAATGTCAACAGAGACTGTGAGAGCATCTCGCGGAGATTTTGTGTCTGAACCCAG	943
QY	812	AAAAAGGCTCAGTCTGGAGCAAAATTCAGAAACATCTCTGG	852
Db	950	CTAAACGCTGTACTCTGAGCAAAATCATGAAGACAAATGG	990
RESULT 14			

AAFF44655  
ID AAF44655 standard; cDNA; 1594 BP.

AC AAF44655;  
yy

DT 27-MAR-2001 (first entry)  
 YY

Novel protein kinase CDNA, SEQ ID NO: 35

KW Human; mouse; protein kinase; antilarthritic; antisclerotic; osteopathic;  
immunopressive: cardiac: renal: antinflammatory: antisthmatic:

OS Mus musculus.

PN WO200073469-A2

PD 07-DEC-2000.  
xy

PF 26-MAY-2000; 2000WO-US14842.  
XY

PR 28-MAY-1999; 9905-0136503  
XX XX

PA (SUGEN INC.)  
XX

XX  
FLOREMAN GD, MALCINEZ R, MYLE D, SUDESSAHAN S,

DR P-PSDB; AAB65629.

PT Nucleic acids enc

PT neurodegenerative diseases and/or cancers -

PS Disclosure; Fig 2; 310pp; English

CC The present sequence encodes a novel protein kinase. The nucleic acids  
CC and the protein kinases they encode may be used in the treatment and  
CC diagnosis of diseases associated with inappropriate kinase expression  
CC such as immune-related diseases and disorders, cardiovascular disease,  
CC neurodegenerative diseases and/or cancers. The nucleic acids and  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays. The kinase polypeptides may be used as antigens in the production  
CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
CC and kinase antagonists may also be used to down regulate kinase  
CC expression and activity. Diseases related to kinase expression and  
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune  
CC disorders, complications of organ transplantation, myocardial infarction,  
CC immune disorders, cardiomyopathies, strokes, renal failure,  
CC oxidative-stress related disorders, chronic inflammatory bowel disease,  
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
CC reproductive disorders.

**SQ** Sequence 1594 BP; 376 A; 488 C; 458 G; 272 T; 0 other;

Query Match	11.2%	Score 261.6;	DB 22;	Length 1594;
Best Local Similarity	58.9%	Pred. No. 3.9e-47;		
Matches 450; Conservative	0;	Mismatches 314;	Indels 0;	Gaps 0;

QY 89 ATGTGGGGCCCTATCGGGCTGGAGAAAGCGCTGGGCAAGACACAGCGGGCTGGTTAAAC 148  
 Db 80 ATGTGGGGCACTATAGGCTGGTTAAGGACCATCGGGAAGGGGCACCTTCGCCAAAATCAAGC 139  
 QY 149 TCGGGGTCCACTGTCATACGAGGTCAGAAAGGTCGCATCAAGATCGTGTAAACCGGAGAAAGC 208  
 Db 140 TGGCTCGGCAATCTCATCGGGGCGGGAGGTCGTTAATGAATCATTTGATTAAGACCAGC 199  
 QY 209 TGTTCGAGTGGTCTGATGAAGGTGGAGCGGGAGATCGCCATCTGTAAAGTCATCGAAC 268

Db 200 TGAACCCAGTAGCTGACAGAGCTGTTACAGAGAAAGTCCGAATTTATGAAGGACTCAACC 259  
 QY 269 ACCGACATGTCCTCAACCTCCAGACGCTTACAGAGAACAGAAATATTTGACGTGTTCC 328  
 Db 260 ACCGACATGTCCTCAACCTTTTGTAGATGATAGACAGCAAGACGCTATCTGTTGA 319  
 QY 329 TGGAGCAGCTGCGGGGGGTGAGCTATTCGACTACCTGCTGTAAGAAAGGAGACAGCC 388  
 Db 320 TGGAAATACGCTAGCGGCAAGAGTGTGACTACCTGCTGCTGCGACGCGCATGAAGG 379  
 QY 389 CCAAGAGGCGCCGAAAGTCTTCCCGCAGATGTGTGCTGCGCTGACCTTCCACAGCT 448  
 Db 380 AGAAGAGGCTCGAGCAAGTTCGCGAGATGCTGACGCGCTGACACTATCTATCAGA 439  
 QY 449 ACTCCATCTCCACAGAGACCTTAAGCCGAGAACTGCTTTGGATGAGAAAAACA 508  
 Db 440 AGAACATTTGACACAGGATCTTAAGGCTGAAGAACTGCTTGTGATCCGAGCCACA 499  
 QY 509 TCCGATTTGACAGATTCGGGATGCGCTCCGACAGTGGGGGACAGCTCCTCGAGACCA 568  
 Db 500 TCAAAATCCGCGACTTCGCTTCAAGATGATGACGCTGCGCTCCAGCTGGACACT 559  
 QY 569 GCTGCGGCTCCCGCATTAATGCTGCTCAGAGTGAATTAAGGGGAAATATGATGACC 628  
 Db 560 TCTGTGGAGACCCCGCATACGCGCGCCAGAGTGTTCAGAGGCAAGATATGATGGCC 619  
 QY 629 GCGGCGACAGATGAGAGCTGTGATCATCTCTTCCCTGCTGCGGGGCTGTC 688  
 Db 620 CAGAGTGGACATCTGAGAGCTGGGTGCTCATCTGTCACCTGTCAGCGGCTCCCTGC 679  
 QY 689 CTTTATGACGACACCTCCGCGCTGCTGAGAGAGTGAAGGAGGAGGCTCTCCACA 748  
 Db 680 CTTTATGAGGACACCTTCAAGAGCTGCGGAGCGAGTCTTCAAGAGAAATGACCGGG 739  
 QY 749 TCGCCCATTTCAATTCCTCAAGATTCAGAGCTCTGAGGAGGATGATCGAAGTGAGC 808  
 Db 740 TCCCTTCTCATGCTGACAGACTGCGAGAGCATTTGCGAGATTTCTGTGCTGAAC 799  
 QY 809 CCGAAAAAGGCTCAGTCTGAGCAATTCAGAAACATCTTGG 852  
 Db 800 CCGCAAAACGCTGTACTGTGAGCAATTCATGAAGACAATGG 843

## RESULT 15

ABA05737 standard; cDNA; 3170 BP.

ABA05737;

04-MAR-2002 (first entry)

Mouse neuronal serine threonine protein kinase 9B5 coding sequence.  
 Mouse; neuronal serine threonine protein kinase; apoplexy; cytosolic;  
 cerebrioprotective; neuroprotective; antiparkinsonian; immunosuppressive;  
 virulence; anti-HIV; antileukemic; antiinflammatory; antitumor; cancer;  
 cerebral infarction; neurological disease; immunological disease; atopy;  
 infection; leukaemia; polyarthritides; Crohn's disease; ulcerative colitis;  
 9B5; ss.

Mus musculus.

Key Location/Qualifiers

1..2175

polyA\_signal

WO20018108-A1.

22-NOV-2001.

XX

PF 17-MAY-2001; 2001MO-EP05660.  
 XX 17-MAY-2000; 2000DE-1024171.  
 PR (BADI) BASF-LYNX BIOSCIENCE AG.  
 PA Schneider A, Klausner B, Fischer A, Newrzella D, Goetz B;  
 PI Rosner M, Eisenhardt G, Kuer R, Trutzel A, Kammandel B;  
 PI Jomana Naim S, Schwaninger M;  
 DR WPI: 2002-055696/07.  
 DR P-PSDB: ABB04431.  
 XX  
 PT New polynucleotide, useful for the diagnosis, treatment and prevention  
 of e.g. apoplexy, tumors and autoimmune disease, comprises a  
 polynucleotide encoding the neuronal protein kinase  
 PS Claim 1; Page 52-53; 75pp; German.  
 CC The present invention provides the protein and coding sequences of two  
 CC neuronal serine threonine protein kinases each from the mouse and human.  
 CC These are designated 9B5 and 9B5.b. Detection of their expression levels  
 CC can be used to diagnose the risk of apoplexy and in prognosis of cerebral  
 CC infarction. They can also be used to identify agents useful in the  
 CC prevention, treatment and diagnosis of neurological diseases, including  
 CC stroke, multiple sclerosis, Parkinson's disease, amyotrophic lateral  
 CC sclerosis, tumours, especially carcinoma, immunological disorders,  
 CC including autoimmune diseases, atopy, viral (including human immune  
 CC deficiency virus) infections, acute or chronic leukocytic or myeloid  
 CC leukaemia, primary chronic polyarthritides, Crohn's disease and ulcerative  
 CC colitis. The present sequence is the murine 9B5 coding sequence.  
 SQ Sequence 3170 BP; 714 A; 988 C; 897 G; 571 T; 0 other;  
 Query Match 11.1%; Score 259.4; DB 24; Length 3170;  
 Best Local Similarity 57.4%; Pred. No. 1.4e-46;  
 Matches 467; Conservative 0; Mismatches 346; Indels 0; Gaps 0;  
 QY 89 ATGTGGCCCTTATGCTGAGAGAGAGCTGGCAAGAGACAGACAGGCTGTTAAAC 148  
 Db 80 ATGTGGCAACTATGAGCTGCTAAGACACATGCGGAGAGGCAACTGCGCAAAAGTCAAGC 139  
 QY 149 TCGGGGTCCATCATCAGAGGCTCAGAAAGTCCGATCAGATGCTGACCGGAGAAC 208  
 Db 140 TGGCTGGCATATCTCAGCGGCGCGGAGGTGCTATTAAGATATGATTAAGACCCAGC 199  
 QY 209 TGTGGAGTGGTGTGATGAGAGTGGAGGAGATCGCATCTGATGAGTCAATGCAAC 268  
 Db 200 TGAACCCAGTAGCTGACAGAACTGTTCAGAGAAAGTCCGAATTTATGAAGGACTCAACC 259  
 QY 269 ACCACATGCTCTCAAGCTCCAGACGCTTACGAGAAACAAGAAATATTTGATCTGTTCC 328  
 Db 260 ACCCAACATCTGGAAGCTTTTGAAGTATGAGAGGAGAAAGACCTATACCTGGTGA 319  
 QY 329 TGGAGCAGCTTCGCGGGGGTACCTTACATACCTGTTAAAGAGGAGGAGTACAGC 388  
 Db 320 TGGAAATACGCTAGCGGAGGAGAGTGTGACTACCTGCTGCGACGCGCATGAAGG 379  
 QY 389 CCAAGAGGCGCCGAAAGTCTTCCCGCAGATGTGTGCTGCGCTGACCTTCCACAGCT 448  
 Db 380 AGAAGAGGCTCGAGCAAGTTCGCGAGATGCTGACGCGCTGACACTATCTATCAGA 439  
 QY 449 ACTCCATCTCCACAGAGACCTTAAGCCGAGAACTGCTTTGGATGAGAAAAACA 508  
 Db 440 AGAACATTTGACACAGGATCTTAAGGCTGAAGAACTGCTTGTGATCCGAGCCACA 499  
 QY 509 TCCGATTTGACAGATTCGGGATGCGCTCCGACAGTGGGGGACAGCTCCTCGAGACCA 568  
 Db 500 TCAAAATCCGCGACTTCGCTTCAAGATGATGACGCTGCGCTCCAGCTGGACACT 559  
 QY 569 GCTGCGGCTCCCGCATTAATGCTGCTCAGAGTGAATTAAGGGGAAATATGATGAGCC 628  
 Db 560 TCTGTGGAGACCCCGCATACGCGCGCCAGAGTGTTCAGAGGCAAGATATGATGGCC 619

QY 629 GCCGGCAGACATGTGAGCTGTGAGTCATCCTTCGCCCTGCTGCGGGCTTC 688  
DB 620 CAGAGGTGACATCTGGAGCCTGGGTGTCACTGTACAGCTGTGTCAGCGGCTCCCTGC 679  
QY 689 CCTTGTATGACGACAACTCCGCCAGCTGCTGGAGAAAGTGAAACGGGCGCTTCCACA 748  
DB 680 CCTTCGATGGGCACAACTCAGAGAGCTGGCGGAGCAATCCTCAGAGAAAGTACCGGG 739  
QY 749 TGCCCCACTTCATTCCTCCAGATTGCCAGAGCCTCTGAGGGGAATGATCGAAGTGAGC 808  
DB 740 TCCCTTCTACATGTCTACAGACTGCGAGAGCATTTCTGCGAGATTCTGTGCTGAACC 799  
QY 809 CCGAAAAAGGCTCAGTCTGAGCAATTCAGAAACATCCTTGTACCTAGCGGGAAAC 868  
DB 800 CCGCAAAAGCTGTACTCTGAGCAATTCAGAAACAAATGATCAACATCGGCTATG 859  
QY 869 ACGAGCCAGACCCGCTGCTGAGCCAGCCCTG 901  
DB 860 AGGGTGAGAGCTGAAGCCATACACGAGGCTG 892

Search completed: April 21, 2003, 20:31:33  
Job time : 337.629 secs

GenCore version 5.1.4\_P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:25:40 ; Search time 3982.82 Seconds

(without alignments)  
17054.732 Million cell updates/sec

Title: US-10-003-690-3

Perfect score: 2334  
Sequence: 1 atgtctccgggggccaaga.....ccaacgggaccctctgccc 2334

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenEmbl:*
1:  gb_ba:*
2:  gb_hlg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

```

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2334	100.0	3109	AF479827	AF479827 Homo sapi
2	2256	96.7	2385	AX166526	AX166526 Sequence
3	2256	96.7	3007	AF479826	AF479826 Homo sapi
4	2223	95.2	2720	HSMB05307	AL834275 Homo sapi
5	2146	91.9	2576	AB058714	AB058714 Homo sapi
6	1470.6	63.0	2128	BC016681	BC016681 Homo sapi
7	961	41.2	2025	AX327995	AX327995 Sequence
8	961	41.2	2219	AX327993	AX327993 Sequence
9	818	35.0	1956	HSAB6701	AX006701 Homo sapi
10	538	23.1	1014	HSMB03233	AL831945 Homo sapi
11	491.8	21.1	3156	AK074411	AK074411 Homo sapi
12	446	19.1	3933	AB014885	AB014885 Halocynthia
13	391.2	16.8	3059	AF316542	AF316542 Caenorhab
14	373.6	16.0	5609	AY060288	AY060288 Drosophila
15	371	15.9	38000	AC008974	AC008974 Homo sapi
16	371	15.9	134792	AC020922	AC020922 Homo sapi
17	315.2	13.5	22606	AC079583	AC079583 Mus muscu
18	314	13.5	210105	AC125948	AC125948 Rattus no
19	307.2	13.2	1873	BC024291	BC024291 Homo sapi
20	265.8	12.7	2333	AF020089	AF020089 Homo sapi
21	265.8	11.4	3529	AB088047	AB088047 Homo sapi
22	265.8	11.4	3609	AB049127	AB049127 Homo sapi
23	264.2	11.3	2462	AX399987	AX399987 Sequence
24	264.2	11.3	3226	AY057448	AY057448 Homo sapi
25	264.2	11.3	3312	AX305105	AX305105 Sequence
26	264.2	11.3	3392	AX305106	AX305106 Sequence
27	264.2	11.3	4917	AB058763	AB058763 Homo sapi
28	261.6	11.2	1594	AX056390	AX056390 Sequence
29	259.4	11.1	3170	AX305103	AX305103 Sequence
30	259.4	11.1	3250	AX305104	AX305104 Sequence
31	245	10.5	2650	AB020480	AB020480 Rattus no
32	245	10.5	4092	AF106937	AF106937 Rattus no
33	240.6	10.3	2112	AX375128	AX375128 Sequence
34	240.6	10.3	2112	AF387638	AF387638 Homo sapi
35	240.6	10.3	2222	AX375131	AX375131 Sequence
36	240.6	10.3	2701	BC008771	BC008771 Homo sapi
37	240.6	10.3	2946	9 HSTPKEWK	X97630 H. sapiens m
38	238.4	10.2	2572	10 RNMARKE2	283869 R. norvegicu
39	237.2	10.2	1749	3 AF020310	AF020310 Drosophila
40	228.8	9.8	2763	10 MKKEM	X70764 M. musculus
41	227.8	9.8	2681	3 AF181649	AF181649 Drosophila
42	225.4	9.7	1869	8 ATRKIN2	X94757 A. thaliana
43	225.4	9.7	2209	8 AY093170	AY093170 Arabidops
44	224.2	9.6	4498	10 MM011494	U11494 Mus musculu
45	221.4	9.5	2352	6 AX262512	AX262512 Sequence

# ALIGNMENTS

RESULT 1  
LOCUS AF479827 3109 bp mRNA linear PRI 12-MAR-2002  
DEFINITION Homo sapiens protein kinase-like protein mRNA, complete cds.  
ACCESSION AF479827  
VERSION AF479827.1 GI:19401873  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 3109)  
AUTHORS She,X.Y., Yu,L. and Guo,J.H.  
TITLE Direct Submission  
JOURNAL Submitted (02-FEB-2002) School of Life Sciences, Laboratory of

Pred. No. is the number of results predicted by chance to have a



```
|||||
Db 1958 CTGGGCTCCCTCCCTTCCACCGGCGAAGTGCAGGTCCTACCGCTGAGAGATGTC 2017
QY 1741 AGCTTGACGCCAGAGTCCCTCCCGAGCTGGCAAAACGCTCTGGTTGGGAATCTATC 1800
Db 2018 AGCTTGACGCCAGAGTCCCTCCCGAGCTGGCAAAACGCTCTGGTTGGGAATCTATC 2077
QY 1801 TCCTTGACAAAGAAAGAAATATCTCTGCTTAAAGACAAACCTCTCAGCAGCATC 1860
Db 2078 TCCTTGACAAAGAAAGAAATATCTCTGCTTAAAGACAAACCTCTCAGCAGCATC 2137
QY 1861 AAGACAGACATCGCATCGCTTCTGTCATGCCACGCTGAGTCAGAGTGTGTCA 1920
Db 2138 AAGACAGACATCGCTTCTGTCATGCCACGCTGAGTCAGAGTGTGTCA 2197
QY 1921 CAGACAGCTTTCAGAGGCGGAGTACAGGCGCAGTGGGCGCCCTCCGTCCTCCAAAAGCC 1980
Db 2198 CAGACAGCTTTCAGAGGCGGAGTACAGGCGCAGTGGGCGCCCTCCGTCCTCCAAAAGCC 2257
QY 1981 GTCCGCTTCCAGTGGAGCATCAGCTCTCTGAGGGTCCAGAGCCCTCCCGGACGGGAC 2040
Db 2258 GTCCGCTTCCAGTGGAGCATCAGCTCTCTGAGGGTCCAGAGCCCTCCCGGACGGGAC 2317
QY 2041 GGCAGCGAGGTGGTGGCATCTACTCCGTCACCTTCACTCATCTGAGGTCCAGCGGT 2100
Db 2318 GGCAGCGAGGTGGTGGCATCTACTCCGTCACCTTCACTCATCTGAGGTCCAGCGGT 2377
QY 2101 CGGTTCAAGCAGAGTGTGGAGACCATCCAGGACAGCTCTGAGCATTATGACAGCC 2160
Db 2378 CGGTTCAAGCAGAGTGTGGAGACCATCCAGGACAGCTCTGAGCATTATGACAGCC 2437
QY 2161 TCCTTGAGAGCCCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
Db 2438 TCCTTGAGAGCCCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2497
QY 2221 CCCGAAGCCTTGCAGCCGCCACCGCGCGCCAGACCCAGAGCTGAGAGCTCTCCCGC 2280
Db 2498 CCCGAAGCCTTGCAGCCGCCACCGCGCGCCAGACCCAGAGCTGAGAGCTCTCCCGC 2557
QY 2281 CGAGGCCGCCGCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2334
Db 2558 CGAGGCCGCCGCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2611

RESULT 2
AX166526 2385 bp DNA linear PAT 22-JUN-2001
LOCUS AX166526
DEFINITION Sequence 17 from Patent WO0138503.
ACCESSION AX166526
VERSION AX166526.1 GI:14546871
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2385)
AUTHORS Plovanin,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
Flanagan,P. and Clardy,D.S.
TITLE Novel human protein kinases and protein kinase-like enzymes
JOURNAL Patent: WO 0138503-A 17 31-MAY-2001;
FEATURES
source location/Qualifiers
1..2385
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 494 a 768 c 731 g 392 t
ORIGIN
Query Match 96.7%; Score 2256; DB 6; Length 2385;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 79 CAGGCCAATATGTGGCCCTATCGGTGAGAGAGAGCTGGGCAAGAGACAGAGAGG 138
|||||
Db 127 CAGGCCAATATGTGGCCCTATCGGTGAGAGAGAGCTGGGCAAGAGACAGAGAGG 186
QY 139 CTGGTTAACTCGGGGTCCTACATGATCAGAGGTGAGAGGTGAGCATCAAGATCTGTAAC 198
Db 187 CTGGTTAACTCGGGGTCCTACATGATCAGAGGTGAGAGGTGAGCATCAAGATCTGTAAC 246
QY 199 CGGAGAGAGTGTGGAGTGGTGTGATGAGAGTGGAGCGGAGAGATCCATCTGTAAG 258
Db 247 CGGAGAGAGTGTGGAGTGGTGTGATGAGAGTGGAGCGGAGAGATCCATCTGTAAG 306
QY 259 CTATGAAACACCATATGTCCTCAAGCTCCAGACGCTACAGAGAAAGAAATATTTG 318
Db 307 CTATGAAACACCATATGTCCTCAAGCTCCAGACGCTACAGAGAAAGAAATATTTG 366
QY 319 TACCTGTTTGGAGACAGCTCTCGGGGGGAGAGTATTCAGCTACCTGTTAAAGAGGG 378
Db 367 TACCTGTTTGGAGACAGCTCTCGGGGGGAGAGTATTCAGCTACCTGTTAAAGAGGG 426
QY 379 AGACTAGCCCAAGAGAGAGCCGAAAGTTCTCCGACAGATTGTCTGGCTGGACTTC 438
Db 427 AGACTAGCCCAAGAGAGAGCCGAAAGTTCTCCGACAGATTGTCTGGCTGGACTTC 486
QY 439 TGGCAGAGTACTCCATCTGTCAGAGACCTAAAGCCGAGAACCTTTTGGATGAG 498
Db 487 TGGCAGAGTACTCCATCTGTCAGAGACCTAAAGCCGAGAACCTTTTGGATGAG 546
QY 499 AAAAACAATCCGATTCGATGAGATTCGGATTCGCTCCGAGTGGGGAGACAGCTC 558
Db 547 AAAAACAATCCGATTCGATGAGATTCGGATTCGCTCCGAGTGGGGAGACAGCTC 606
QY 559 CTGAGACAGAGTGGCGGGTCCCGCAATTATGCTGTCCAGAGTGTATTAAAGGGGAAAA 618
Db 607 CTGAGACAGAGTGGCGGGTCCCGCAATTATGCTGTCCAGAGTGTATTAAAGGGGAAAA 666
QY 619 TATGATGGCGCGGGGAGAGATGTGAGAGTGTGAGATCTCTTGGCTGTCTG 678
Db 667 TATGATGGCGCGGGGAGAGATGTGAGAGTGTGAGATCTCTTGGCTGTCTG 726
QY 679 GGGGCTCTGCCCTTGTATGAGAGCAACCTCCGCGAGCTGTGAGAGAGTAAAGGGGGC 738
Db 727 GGGGCTCTGCCCTTGTATGAGAGCAACCTCCGCGAGCTGTGAGAGAGTAAAGGGGGC 786
QY 739 GTCTTCCATATGCCCATCTTCATCTCCAGATTCAGAGACCTCTCTGAGGGGAATATC 798
Db 787 GTCTTCCATATGCCCATCTTCATCTCCAGATTCAGAGACCTCTCTGAGGGGAATATC 846
QY 799 GAGGTGAGCGCGGAAAAAGGCTGAGTGAAGCAAAATTCAGAAACATCTTGTGACCTA 858
Db 847 GAGGTGAGCGCGGAAAAAGGCTGAGTGAAGCAAAATTCAGAAACATCTTGTGACCTA 906
QY 859 GGGGGAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 918
Db 907 GGGGGAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 966
QY 919 CGAGAGCTGCTCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 978
Db 967 CGAGAGCTGCTCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1026
QY 979 GGGTCTTCAAGGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1038
Db 1027 GGGTCTTCAAGGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1086
QY 1039 GAAAGATGATATATATATCTCTTTTGGATCGAAGAGAGAGAGAGAGAGAGAGAGAG 1098
Db 1087 GAAAGATGATATATATATCTCTTTTGGATCGAAGAGAGAGAGAGAGAGAGAGAGAG 1146
QY 1099 CAGGAGCTGCTCCCGGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1158
Db 1147 CAGGAGCTGCTCCCGGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1206
QY 1159 CTGAGCGGTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1218
```

Db 1207 CTGAGCCGTCACGGGAAAGCGGACGACAGAGCGGAATCCATGGAAGTCCCTGAGCATCAC 1266  
 QY 1219 GATGCCGGGGGGTGGTGGCTCCCTGTACACACCGAGGGGCTTGGAGATGGCCAGAC 1278  
 Db 1267 GATGCCGGGGGGTGGTGGCTCCCTGTACACACCGAGGGGCTTGGAGATGGCCAGAC 1326  
 QY 1279 AGCCAGAGATCCGTAGCGTCAATGAGACCTTCACAGGGTCTGTCTCCACAGCCCTCTAAC 1338  
 Db 1327 AGCCAGAGATCCGTAGCGTCAATGAGACCTTCACAGGGTCTGTCTCCACAGCCCTCTAAC 1386  
 QY 1339 AGCCAGAGATCCGTAGCGTCAATGAGACCTTCACAGGGTCTGTCTCCACAGGGTCTGTAC 1398  
 Db 1387 AGCCAGAGATCCGTAGCGTCAATGAGACCTTCACAGGGTCTGTCTCCACAGGGTCTGTAC 1446  
 QY 1399 GGGGGGGGGTCCCGGATCCCAAAAGCAGAGCGTCTCTCGGGGGCCAGGGGGTGG 1458  
 Db 1447 GGGGGGGGGTCCCGGATCCCAAAAGCAGAGCGTCTCTCGGGGGCCAGGGGGTGG 1506  
 QY 1459 GGGGGGGGGAGAGCCCGCCCGCCAGAGTCCCGCTCCACACCCCTGCGGGCCCGCA 1518  
 Db 1507 GGGGGGGGGAGAGCCCGCCCGCCAGAGTCCCGCTCCACACCCCTGCGGGCCCGCA 1566  
 QY 1519 GGGTCCCGGGCTCTCTGGGGGGAGCCCTTGGACCTGCTCTGACAGAGCCCGGGGCG 1578  
 Db 1567 GGGTCCCGGGCTCTCTGGGGGGAGCCCTTGGACCTGCTCTGACAGAGCCCGGGGCG 1626  
 QY 1579 AGTCCACCGGAGACCCCGGGGACAAACACACCCCGGCGGGTGGCGTGGGGGA 1638  
 Db 1627 AGTCCACCGGAGACCCCGGGGACAAACACACCCCGGCGGGTGGCGTGGGGGA 1686  
 QY 1639 GCGCGCTGGAGAGTGTCTCACTCCAGCCGCAACAGCTTCTGGGCTCCCTGCTTT 1698  
 Db 1687 GCGCGCTGGAGAGTGTCTCACTCCAGCCGCAACAGCTTCTGGGCTCCCTGCTTT 1746  
 QY 1699 CACCGGGCAAGATGAGGTCTACCGCTGAGGAGATGTCAGTTCAGCGCAGAGTCC 1758  
 Db 1747 CACCGGGCAAGATGAGGTCTACCGCTGAGGAGATGTCAGTTCAGCGCAGAGTCC 1806  
 QY 1759 TCCCGGAGCTGGCAAAAGCGTCTGTTGCGGAATCTCATCTCTGGAGCAAAAGAA 1818  
 Db 1807 TCCCGGAGCTGGCAAAAGCGTCTGTTGCGGAATCTCATCTCTGGAGCAAAAGAA 1866  
 QY 1819 CAATATCTCTCTGCTGCTAAAGAGCAAACTCTCAGCAGCATCAAGCAGATCTCCAT 1878  
 Db 1867 CAATATCTCTCTGCTGCTAAAGAGCAAACTCTCAGCAGCATCAAGCAGATCTCCAT 1926  
 QY 1879 GCTTTCTGTCGATCCCGAGCTGATCAGAGTGTCTGACAGACAGCTTCAGGGGC 1938  
 Db 1927 GCTTTCTGTCGATCCCGAGCTGATCAGAGTGTCTGACAGACAGCTTCAGGGGC 1986  
 QY 1939 GAGTACAAAGCCAGTGGCGGCGCTCGCTTCCAAAAGCCGCTCCGCTCCAGGTGAC 1998  
 Db 1987 GAGTACAAAGCCAGTGGCGGCGCTCGCTTCCAAAAGCCGCTCCGCTCCAGGTGAC 2046  
 QY 1999 ATCAGCTCTCTGAGGTTCAGAGCCCTCCCGGAGCGGAGCGGAGGTGGTGC 2058  
 Db 2047 ATCAGCTCTCTGAGGTTCAGAGCCCTCCCGGAGCGGAGCGGAGGTGGTGC 2106  
 QY 2059 ATCTACTCGTCACTCTCATCTCGGGGTCCAGCGGTCCGTTCAAGGAGTGGTG 2118  
 Db 2107 ATCTACTCGTCACTCTCATCTCGGGGTCCAGCGGTCCGTTCAAGGAGTGGTG 2166  
 QY 2119 GAGACATCCAGGACAGTCTCTGAGCACTGATGACAGCGCTCCGTTCAAGGAGTGGTG 2178  
 Db 2167 GAGACATCCAGGACAGTCTCTGAGCACTGATGACAGCGCTCCGTTCAAGGAGTGGTG 2226  
 QY 2179 GACGAGAAGACGGGGGCCAGAGCCCGCTGCTGGTGGCCCAAGGAGCTTCAGGCC 2238  
 Db 2227 GACGAGAAGACGGGGGCCAGAGCCCGCTGCTGGTGGCCCAAGGAGCTTCAGGCC 2286  
 QY 2239 CACCCGGGGCCCGAGACCCAGAGCTGAGCAGCTCTCCCGCGAGGGCCCGCAAGGAC 2298  
 Db 2287 CACCCGGGGCCCGAGACCCAGAGCTGAGCAGCTCTCCCGCGAGGGCCCGCAAGGAC 2346

QY 2299 AAGAAGCTCTGGCCACACAGGAGCCCTCTGCC 2334  
 Db 2347 AAGAAGCTCTGGCCACACAGGAGCCCTCTGCC 2382

RESULT 3  
 AF479826  
 LOCUS  
 DEFINITION Homo sapiens putative serine/threonine protein kinase mRNA,  
 complete cds.  
 ACCESSION AF479826  
 VERSION AF479826.1 GI:19401870  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
 1 (bases 1 to 3007)  
 AUTHORS She, X.Y., Guo, J.H. and Yu, L.  
 TITLE Submitted (02-FEB-2002) School of Life Sciences, Laboratory of Human Genes Research, Institute of Genetics, Fudan University, 220 Handan Road, Shanghai 200433, P. R. China  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source  
 1..3007  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /map="19q13"  
 /tissue="brain"  
 128..2512  
 /cfeature="exon"  
 /product="putative serine/threonine protein kinase"  
 /protein\_id="AA187697.1"  
 /db\_xref="GI:19401871"  
 /translation="MVAAGLTLGSPSPDQVSPERKDEAVAGGEGEERAEERGRHA

## CDS

QY 79 CACGCCAATATGTGGGCGCTATCGGCTGGAGAAGCGTGGGCAAGGACAGACAGG 138  
 Db 254 CACGCCAATATGTGGGCGCTATCGGCTGGAGAAGCGTGGGCAAGGACAGACAGG 313  
 QY 139 CTGGTTAACTCGGGGTCCACTCAGCAGGAGTCAAGAGTCCGCAATCAAGATGTGAAC 198  
 Db 314 CTGGTTAACTCGGGGTCCACTCAGCAGGAGTCAAGAGTCCGCAATCAAGATGTGAAC 373  
 QY 199 CGGAGAGAGCTGTGAGTGGTGTGATGAGGTGAGCGGGAGATCGCCATCTGTAAG 258  
 Db 374 CGGAGAGAGCTGTGAGTGGTGTGATGAGGTGAGCGGGAGATCGCCATCTGTAAG 433  
 QY 259 CTCATGAGACACCAATGCTCTCAAGCTCCAGAGCTCTACAGGAACAAGAAATATTG 318  
 Db 434 CTCATGAGACACCAATGCTCTCAAGCTCCAGAGCTCTACAGGAACAAGAAATATTG 493

Query Match 96.7%; Score 2256; DB 9; Length 3007;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



REFERENCE 1 (bases 1 to 2720)  
 Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.  
 Koehner, K., Beyer, A., Mewes, H.W., Well, B. and Wiemann, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-JUL-2002) 1. D-85764 Neuherberg, GERMANY  
 Clome from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;  
 Berlin-Charlottenburg, GERMANY; Email: clone@zpd.de Further  
 Berlin/Germany) within the CDNA sequencing consortium of the German  
 genome project.  
 This clone (DKFZp547E1613) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@zpd.de Further  
 information about the clone and the sequencing project is available  
 at http://mhps.gsf.de/proj/cDNA/.

## FEATURES

source  
 1. 2720  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /clone="DKFZp547E1613"  
 /tissue\_type="brain"  
 /clone\_lib="547 (synonym: hfbri). Vector pSport1; host  
 DH10B; sites NotI + SalI"  
 /dev\_stage="fetal"  
 1. 2720  
 /gene="DKFZp547E1613"  
 /note="putative serine/threonine protein kinase"  
 /product="hypothetical protein"  
 /protein\_id="CAD38950.1"  
 /db\_xref="GI:21739847"  
 /translation="RLEKTLGKGTGLVKGHCITGOKRAIKIVNEKLSSTLMV  
 EREAIKRIEHPHYLKLVDYENKRYLYLVHVSSEGLPDLVKGRIPLKARKF  
 FQIVSADFCSTYSICHRDLKPENLLEKNNIRADFMASLQVDSLETSCGSP  
 HYACEVINGEKYDGRADMSGVILFALVGLPDDNLSLQLEKVRVARSIP  
 FLPPDQSLRLGMIWEPEKRLSLQIHKHPVYLGKHEDEPCLEPARGVARSIP  
 SGEIDPVLSEMASLIGCFRDRRLHRELESEEDNKMYLLLDKREYPSCEDD  
 LEPNDVDPKRRVDSPLSRHGKRRPERKSMELSTTDGSGSPPTRALEMAH  
 SORRSVSASTGSSPLSSPSRSPYSPSPSPSPSPSPSPSPSPSPSPSPSPSP  
 GGGAGEQPPPPARSTPLPGPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP  
 GGGAGAMRSLNLSIRNSFLGSPFRHKKQVPAEHSLSLTPSSPELAKRSGVNF  
 SLDKKEQJFLVLEKSLKADIVHAFISIPSLSVLSQTSFRAFKASGCVSVQ  
 KVRVQVDSISSEGGPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP  
 HDQSVQALADEKNGKQTRPAGAPPSLQPPPPRPPDELLSSFRGPPQGEARHGR  
 DPSALTFRGRGRGPPSTPLPCPTVNL"  
 2703  
 /gene="DKFZp547E1613"  
 /note="putative serine/threonine protein kinase"  
 /product="hypothetical protein"  
 /protein\_id="CAD38950.1"  
 /db\_xref="GI:21739847"  
 /translation="RLEKTLGKGTGLVKGHCITGOKRAIKIVNEKLSSTLMV  
 EREAIKRIEHPHYLKLVDYENKRYLYLVHVSSEGLPDLVKGRIPLKARKF  
 FQIVSADFCSTYSICHRDLKPENLLEKNNIRADFMASLQVDSLETSCGSP  
 HYACEVINGEKYDGRADMSGVILFALVGLPDDNLSLQLEKVRVARSIP  
 FLPPDQSLRLGMIWEPEKRLSLQIHKHPVYLGKHEDEPCLEPARGVARSIP  
 SGEIDPVLSEMASLIGCFRDRRLHRELESEEDNKMYLLLDKREYPSCEDD  
 LEPNDVDPKRRVDSPLSRHGKRRPERKSMELSTTDGSGSPPTRALEMAH  
 SORRSVSASTGSSPLSSPSRSPYSPSPSPSPSPSPSPSPSPSPSPSPSPSP  
 GGGAGEQPPPPARSTPLPGPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP  
 GGGAGAMRSLNLSIRNSFLGSPFRHKKQVPAEHSLSLTPSSPELAKRSGVNF  
 SLDKKEQJFLVLEKSLKADIVHAFISIPSLSVLSQTSFRAFKASGCVSVQ  
 KVRVQVDSISSEGGPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP  
 HDQSVQALADEKNGKQTRPAGAPPSLQPPPPRPPDELLSSFRGPPQGEARHGR  
 DPSALTFRGRGRGPPSTPLPCPTVNL"

polyA\_site  
 BASE COUNT 551 a 879 c 807 g 483 t  
 ORIGIN

Query Match 95.2%; Score 2223; DB 9; Length 2720;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2234; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 101 ATCGGCTGAGAAAGAGCTGGGCAAGAGACAGAGGCTGGTTAACTGGGGTCCACT 160  
 DB 1 ATCGGCTGAGAAAGAGCTGGGCAAGAGACAGAGGCTGGTTAACTGGGGTCCACT 60  
 QY 161 GCATCAGGGGTGAGAGGTTCGCATCAAGATCGTGAACCGGGAGAAAGCTGCGAGTCGG 220  
 DB 61 GCATCAGGGGTGAGAGGTTCGCATCAAGATCGTGAACCGGGAGAAAGCTGCGAGTCGG 120  
 QY 221 TGCCTGATGAAGTGGAGCGGGAGATGCGCATCTCTGAAGCTCATCGAACACCCACATGTC 280  
 DB 121 TGCCTGATGAAGTGGAGCGGGAGATGCGCATCTCTGAAGCTCATCGAACACCCACATGTC 180  
 QY 281 TCAAGCTCAGACGCTCTAGAGAACAGAAATATTGTACTGCTGCTGAGAGACGCT 340  
 DB 181 TCAAGCTCAGACGCTCTAGAGAACAGAAATATTGTACTGCTGCTGAGAGACGCT 240

QY 341 CGGGGGGTGAGCTATTGCACTACCTGTGTAAGAGAGGAGACTGACGCCAAGAGGCC 400  
 DB 241 CGGGGGGTGAGCTATTGCACTACCTGTGTAAGAGAGGAGACTGACGCCAAGAGGCC 300  
 QY 401 GAAAGTCTTCGCGCAGATTTGTCTGCGCTGAGCTTCTCCACAGCTACTCATCTGCC 460  
 DB 301 GAAAGTCTTCGCGCAGATTTGTCTGCGCTGAGCTTCTCCACAGCTACTCATCTGCC 360  
 QY 461 ACAGAGACCTTAAGCCCGAGAACCTGCTTGGATGAGAAACAAATCCGCTTTCAG 520  
 DB 361 ACAGAGACCTTAAGCCCGAGAACCTGCTTGGATGAGAAACAAATCCGCTTTCAG 420  
 QY 521 ACTTCGCAATGAGCGTCCCTGCAAGTGGGAGACACCTCTGAGACAGCTGCGGTCGC 580  
 DB 421 ACTTCGCAATGAGCGTCCCTGCAAGTGGGAGACACCTCTGAGACAGCTGCGGTCGC 480  
 QY 581 CCCATTATGCGTGTCCAGAGGTGATTAAAGGGGAAAAAATATGATGCGCGCGGAGACA 640  
 DB 481 CCCATTATGCGTGTCCAGAGGTGATTAAAGGGGAAAAAATATGATGCGCGCGGAGACA 540  
 QY 641 TGTGAGAGCTGTGAGATCATTCTTCCCTGCTGCTGCGGGGGCTCTGCTTGTATGACG 700  
 DB 541 TGTGAGAGCTGTGAGATCATTCTTCCCTGCTGCTGCGGGGGCTCTGCTTGTATGACG 600  
 QY 701 ACAACCTCCGCGCAGCTGTGAGAGAGGTGAACGGGGCGTCTTCCACATGCGCCACTTCA 760  
 DB 601 ACAACCTCCGCGCAGCTGTGAGAGAGGTGAACGGGGCGTCTTCCACATGCGCCACTTCA 660  
 QY 761 TTCCTCCAGATTGCCAGAGCTCTCTGAGGGGAATGATCGAAGTGAAGCCCGAAAAAGGC 820  
 DB 661 TTCCTCCAGATTGCCAGAGCTCTCTGAGGGGAATGATCGAAGTGAAGCCCGAAAAAGGC 720  
 QY 821 TCAGTCTGAGCAAAATTCGAAATCATCTTGTGTAAGTGAAGGCGGAAACAGAGCAAGC 880  
 DB 721 TCAGTCTGAGCAAAATTCGAAATCATCTTGTGTAAGTGAAGGCGGAAACAGAGCAAGC 780  
 QY 881 CGTGCCTGAGAGCAGCCCTGCGCGCGCGGAGTACCATGCGAGACCTGCTCCATCCAGAG 940  
 DB 781 CGTGCCTGAGAGCAGCCCTGCGCGCGCGGAGTACCATGCGAGACCTGCTCCATCCAGAG 840  
 QY 941 AACTGAGACCCCGACGCTCTGAGAGCATATGAGCTGAGCTGAGGCTCTTCAAGGACCGGAGA 1000  
 DB 841 AACTGAGACCCCGACGCTCTGAGAGCATATGAGCTGAGCTGAGGCTCTTCAAGGACCGGAGA 900  
 QY 1001 GCGTGCATGCGCGAGCTGCGAGTGAAGAGAGAACCAAGAAAGATATATATATCTGC 1060  
 DB 901 GCGTGCATGCGCGAGCTGCGAGTGAAGAGAGAACCAAGAAAGATATATATATCTGC 960  
 QY 1061 TTTTGGATGGAAGAGAGCGATATCCAGCTGTGAGAGACAGACCTGCTCCCGGAATG 1120  
 DB 961 TTTTGGATGGAAGAGAGCGATATCCAGCTGTGAGAGACAGACCTGCTCCCGGAATG 1020  
 QY 1121 ATGTTGACCCCGCGGAGAGCGTGTGATTTCTCCATGCTGAGCGCTGACGGGAAGCGGC 1180  
 DB 1021 ATGTTGACCCCGCGGAGAGCGTGTGATTTCTCCATGCTGAGCGCTGACGGGAAGCGGC 1080  
 QY 1181 GACCAAGAGGAGAGTCCATGGAAGTCCAGATCCATGACCAAGTCCGGGGGTGTGGCTCCC 1240  
 DB 1081 GACCAAGAGGAGAGTCCATGGAAGTCCAGATCCATGACCAAGTCCGGGGGTGTGGCTCCC 1140  
 QY 1241 CTGTACCCACCCGAGCGGCTTGGAGATGAGCCAGCAGACCTGCTCCGAGAGCTCA 1300  
 DB 1141 CTGTACCCACCCGAGCGGCTTGGAGATGAGCCAGCAGACCTGCTCCGAGAGCTCA 1200  
 QY 1301 GTGAGACCTCCACGGGTCTGTCTCTCCAGCTCTTAAGCAGCCCAAGAGAGTCCGCTTTT 1360  
 DB 1201 GTGAGACCTCCACGGGTCTGTCTCTCCAGCTCTTAAGCAGCCCAAGAGAGTCCGCTTTT 1260  
 QY 1361 CTTTTCACCGAGCGCGGGGGCTGGAGATGAGGCTGAGAGCGGGGGCTCCCGACATTCA 1420  
 DB 1261 CTTTTCACCGAGCGCGGGGGCTGGAGATGAGGCTGAGAGCGGGGGCTCCCGACATTCA 1320



489 TTTGATGAGAAAAACACATCCGATTGACATTCGGATGGGCTGCTTCAGAGTGGG 548  
 301 TTTGATGAGAAAAACACATCCGATTGACATTCGGATGGGCTGCTTCAGAGTGGG 360  
 549 GACAGACCTCTGAGAGACAGCTGGGGTCCCGCATTTATGCTGTCAGAGTATTA 608  
 361 GACAGACCTCTGAGAGACAGCTGGGGTCCCGCATTTATGCTGTCAGAGTATTA 420  
 609 GGGGAAAAATATGATGGCGCGGAGAGATGATGGAGCTGGAGATCCTTTGGC 668  
 421 GGGGAAAAATATGATGGCGCGGAGAGATGATGGAGCTGGAGATCCTTTGGC 480  
 669 CCGCTGCTGGGGCTTGGCTTTGATGAGACAACTCCCGCAGCTGCTGAGAGAT 728  
 481 CCGCTGCTGGGGCTTGGCTTTGATGAGACAACTCCCGCAGCTGCTGAGAGAT 540  
 729 GAAACGGGCGCTTCCACATGCCCACTTCCTTCAGATTCAGAGCTCTGAG 788  
 541 GAAACGGGCGCTTCCACATGCCCACTTCCTTCAGATTCAGAGCTCTGAG 600  
 789 GGAATGATGAGATGAGAGCCGAAAAAGGCTGATGAGAGAAATTCAGAAATCC 848  
 601 GGAATGATGAGATGAGAGCCGAAAAAGGCTGATGAGAGAAATTCAGAAATCC 660  
 849 TTGCTACTAGGCGGAGAAACAGAGACCGCTGCTGAGAGCCGCTGGCGCG 908  
 661 TTGCTACTAGGCGGAGAAACAGAGACCGCTGCTGAGAGCCGCTGGCGCG 720  
 909 GGTAGCATCGGAGACCTGCCATCCAAAGAGAGCTGGAGCCGCTGCTAGAGCAT 968  
 721 GGTAGCATCGGAGACCTGCCATCCAAAGAGAGCTGGAGCCGCTGCTAGAGCAT 780  
 969 GGCATCATCTGGGCTGCTTACAGGAGACCGGAGAGGCTGATGCGGAGTGA 1028  
 781 GGCATCATCTGGGCTGCTTACAGGAGACCGGAGAGGCTGATGCGGAGTGA 840  
 1029 GGAAGAACCAAGAAAGATATATATATCTCTTTGATGAGAGAGAGCGTATCCAG 1088  
 841 GGAAGAACCAAGAAAGATATATATATCTCTTTGATGAGAGAGAGCGTATCCAG 900  
 1089 CTGTGAGAGACGAGACCTGCTCCCGGAAATGATGTTGACCCCGCGGAAAGCTGGA 1148  
 901 CTGTGAGAGACGAGACCTGCTCCCGGAAATGATGTTGACCCCGCGGAAAGCTGGA 960  
 1149 TTCTCCATGCTGAGAGCTGCTCAAGGAGAGCGGAGACAGAGGAGTGCATGAGTCT 1208  
 961 TTCTCCATGCTGAGAGCTGCTCAAGGAGAGCGGAGACAGAGGAGTGCATGAGTCT 1020  
 1209 GAGCATCACGATGCGGGGGTGGTGGCTCCCTGTACCAACCGGAGCGGCTTGGAGAT 1268  
 1021 GAGCATCACGATGCGGGGGTGGTGGCTCCCTGTACCAACCGGAGCGGCTTGGAGAT 1080  
 1269 GGGCCACAGACGAGACGATCCGATGAGAGAGCTGAGAGCGGAGTCTGCTCCAG 1328  
 1081 GGGCCACAGACGAGACGATCCGATGAGAGAGCTGAGAGCGGAGTCTGCTCCAG 1140  
 1329 CCTCTAAGACAGCCCAAGAGATGCTGCTTTTCTTTTCTTACCGAGAGCGGGGCTGAGA 1388  
 1141 CCTCTAAGACAGCCCAAGAGATGCTGCTTTTCTTTTCTTACCGAGAGCGGGGCTGAGA 1200  
 1389 TGAAGCTGAGAGCGGGGCTCCCGACTTCCAAAAAGCAGAGAGCTGCTGCGGGGCC 1448  
 1201 TGAAGCTGAGAGCGGGGCTCCCGACTTCCAAAAAGCAGAGAGCTGCTGCGGGGCC 1260  
 1449 CAGGAGTGGGGGCGCGGAGAGAGCGGCGGCGGCGGCGGAGTCTGCAACCGCTGCG 1508  
 1261 CAGGAGTGGGGGCGCGGAGAGAGCGGCGGCGGCGGCGGAGTCTGCAACCGCTGCG 1320  
 1509 CAGGAGTGGGGGCGCGGAGAGAGCGGCGGCGGCGGCGGAGTCTGCAACCGCTGCG 1568  
 1321 CAGGAGTGGGGGCGCGGAGAGAGCGGCGGCGGCGGAGTCTGCAACCGCTGCG 1380

1569 GCCCGGGGCGAGTCCACCGGAGACCCCGGGGACAAACACACCCCGGCGGCTG 1628  
 1381 GCCCGGGGCGAGTCCACCGGAGACCCCGGGGACAAACACACCCCGGCGGCTG 1440  
 1629 CGTGGGGGAGCGGCTGAGAGAGTCTCTCACTCATCCAGACGCTTCTGGGCTC 1688  
 1441 CGTGGGGGAGCGGCTGAGAGAGTCTCTCACTCATCCAGACGCTTCTGGGCTC 1500  
 1689 CCGTCTTTCACCGGCGGAGAGTCCCTACCGCTGAGAGAGTCTGAGCTTGC 1748  
 1501 CCGTCTTTCACCGGCGGAGAGTCCCTACCGCTGAGAGAGTCTGAGCTTGC 1560  
 1749 GCCAGAGTCTCCCGGAGAGTGGCAAAACCTCTGCTGGGAGTCAATCTCTTGA 1808  
 1561 GCCAGAGTCTCCCGGAGAGTGGCAAAACCTCTGCTGGGAGTCAATCTCTTGA 1620  
 1809 CAAAGAAACAAATATCTCTGCTGCTTAAAGAGACAACTCTGAGCATCAAGAGCA 1868  
 1621 CAAAGAAACAAATATCTCTGCTGCTTAAAGAGACAACTCTGAGCATCAAGAGCA 1680  
 1869 CATGCTCATGCTCTTCTGCTGATCCCGAGCTGAGTACAGTGTGCTGACAGACAG 1928  
 1681 CATGCTCATGCTCTTCTGCTGATCCCGAGCTGAGTACAGTGTGCTGACAGACAG 1740  
 1929 CTTACAGGCGGAGTACAGAGCGAGTGGCGGCGGCTCTGCTTCCAAAAAGCGGCTT 1988  
 1741 CTTACAGGCGGAGTACAGAGCGAGTGGCGGCGGCTCTGCTTCCAAAAAGCGGCTT 1800  
 1989 CCAGGTGAGATCAGTCTCTGAGGTTCCAGAGCTCTCCCGGAGAGGAGCGGAGG 2048  
 1801 CCAGGTGAGATCAGTCTCTGAGGTTCCAGAGCTCTCCCGGAGAGGAGCGGAGG 1860  
 2049 AGGTGAGTATCTACTCTGCTGATCCCTGATCTGATCTGAGGTTCCAGAGCTGCTTCA 2108  
 1861 AGGTGAGTATCTACTCTGCTGATCCCTGATCTGATCTGAGGTTCCAGAGCTGCTTCA 1920  
 2109 GCGAGTGTGAGAGACCTTCCAGAGAGCTCTGAGAGCTCTGAGAGCTCTGAGAG 2168  
 1921 GCGAGTGTGAGAGACCTTCCAGAGAGCTCTGAGAGCTCTGAGAGCTCTGAGAG 1980  
 2169 GGCCTGTGAGAGAGAGAGAGAGAGGCGGCGGAGAGCTCTGAGAGCTCTGAGAG 2228  
 1981 GGCCTGTGAGAGAGAGAGAGAGAGGCGGCGGAGAGCTCTGAGAGCTCTGAGAG 2040  
 2229 CCTGAGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2288  
 2041 CCTGAGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
 2289 CCGCAAG 2334  
 2101 CCGCAAG 2146

RESULT 6  
 BC016681 2128 bp mRNA linear PRI 05-NOV-2001  
 LOCUS Homo sapiens, clone MGC:21870 IMAGE:4341838, mRNA, complete cds.  
 DEFINITION  
 ACCESSION BC016681  
 VERSION BC016681.1 GI:16741775  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 1 (bases 1 to 2128)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-OCT-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 COMMENT Contact: MGC help desk



```

Db 1609 CAGAGAGCTCTGGCCACCAAGGAGCCCTCTGCC 1645
RESULT 7
AX327995
LOCUS AX327995 2025 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 3 from Patent WO0181588.
ACCESSION AX327995
VERSION AX327995.1 GI:18098148
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1. Meyers, R.
2246. Protein Kinase molecules and uses therefor
PATENT: WO 0181588-A 3 01-NOV-2001.
JOURNAL Millennium Pharmaceuticals, Inc. (US)
FEATURES
source location/Qualifiers
1. 2025
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 442 a 641 c 616 g 326 t
ORIGIN
Query Match 41.2%; Score 961; DB 6; Length 2025;
Best Local Similarity 70.1%; Pred. No. 5.8e-164;
Matches 1477; Conservative 0; Mismatches 435; Indels 195; Gaps 6;
QY 74 CCCAGACGCCCAATATGTGGCCCTATGCGCTGGAGAAACGCTGGGCAAGACAGA 133
DB 29 CGAGACACGGCGATGTGGCCCTACCGCTGGAGAAACGCTGGGCAAGACAGA 88
QY 134 CAGGCTGTAACTGGGGTCCATGTCATCAGCGGTCAGAGGTCCGCTCAAGATCG 193
DB 89 CAGGTCTGTGAAGCTGGGGTTCACCTGCTCAGCTGCAAGAGTGGCCATCAAGATCG 148
QY 194 TGAACCGGGAAGAGTGTGCGAGTGTCTGTATGAAGTGGAGCGGAGATCGCCATCC 253
DB 149 TCAACCGTGAAGAGTGTGCGAGTGTCTGTATGAAGTGGAGCGGAGATCGCCATCC 208
QY 254 TGAAGTCAACGAAACACCCATATGCTCTCAAGCTCCACGAGTCTACAGAGAAACAAT 313
DB 209 TGAAGTCAATTTGAGAGTCCACCTCAAGCTCTAAAGTGCACAGCTTTTAAAAAATA 268
QY 314 ATTTGTAAGTGTCTGTGAGACAGTCTCGGGGGGTGAGACTATTGCACTACCTGGTAAAGA 373
DB 269 ATTTGTAAGTGTCTGTGAGACAGTCTCGGGGGGTGAGACTATTGCACTACCTGGTAAAGA 328
QY 374 AGGGAAGACTGAGCGCCCAAGAGAGGCCGAAAGTTCTTCGCGCAAGATTGTCTGCGCTGG 433
DB 329 AGGGAAGACTGAGCGCCCAAGAGAGGCCGAAAGTTCTTCGCGCAAGATTGTCTGCGCTGG 388
QY 434 ACTTTCGCAACAGTCTACTCATCTGCAACAGAGACTAAACCCGAGAAACCTGCTTTGG 493
DB 389 ACTTTCGCAACAGTCTACTCATCTGCAACAGAGACTAAACCTGCTTTGG 448
QY 494 ATGAGAAAAACAACATCCGATTCGACAGATTGCGCATGGCTCCCTCGAGGAGGAGACA 553
DB 449 ATGAGAAAAACAACATCCGATTCGACAGATTGCGCATGGCTCCCTCGAGGAGGAGACA 508
QY 554 GCCTCTGAGACAGTGTGGGTCCGCCATTATGCGTGTCAAGAGTGAATTAAGGGGG 613
DB 509 GCCTGTGAGACAGTGTGGGTCCGCCATTATGCGTGTCAAGAGTGAATTAAGGGGG 568
QY 614 AAAAATATATGAGCGCGCGGAGACATGTGAGACTGTGAGATCATCTTCCGCTGCG 673
DB 569 AGAATATATGAGCGCGCGGAGACATGTGAGACTGTGAGATCATCTTCCGCTGCG 628
QY 674 TGTGGGGGCTGCGCTTGTATGAGACAACCTCCGCACTGCTGTGAGAGTGAAC 733

```

```

Db 629 TGTGGGGGCTTGCCTTGTGAGATGACAACCTTGCACAGCTGTGGAGAGTGAAC 688
QY 734 GGGGCTCTTCCACATGCCCCCACTTCACTTCCAGATTTGCAAGAGCTCTGAGGGAA 793
DB 689 GGGGCTCTTCCACATGCCCCCACTTATCCGCGGAGTCCACAGTCTGTAGGGGCA 748
QY 794 TGTTCGAGTGTGAGCGCCGAAAGAAAGCTCAGTCTGTGAGCAATTTAGAAACATCTTGGT 853
DB 749 TGTTCGAGTGTGAGCGCGGAGAGAGCTCAGCTGTGAGCAATTTAGAAACATATGTGT 808
QY 854 ACCTAGCGGGAACACAGACAGCCGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 913
DB 809 ATATAGGGGCAAGATGATGACCCGAAAC-----AGAGCAAGCCATTCTTCCAGAGTGC 862
QY 914 CCATGGGAGACCTTCCATCCACGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 973
DB 863 AGATCCGCTGCTGCGCAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 922
QY 974 CACTGGGCTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1033
DB 923 CACTGGGCTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 982
QY 1034 ACCAAGAAAAAGATATATATATATCTGCTTTGATGAGAGAGAGAGAGAGAGAGAGAG 1093
DB 983 ACCAAGAAAAAGATATATATATCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1042
QY 1094 AGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1153
DB 1043 AGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1102
QY 1154 CCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1213
DB 1103 CCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1162
QY 1214 TCAACGATGCGGGGAGTGTGCTTCCCTTACACCAAGAGAGAGAGAGAGAGAGAGAG 1273
DB 1163 TCAAC-----GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1213
QY 1274 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1333
DB 1214 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1273
QY 1334 TTAGCAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1393
DB 1274 TTAGCAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1286
QY 1394 CTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1453
DB 1287 ----- 1286
QY 1454 GTGGGGGCGCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1513
DB 1287 -----GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1326
QY 1514 CCCCAGAGCTCCCGGCGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1573
DB 1327 -----ACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1357
QY 1574 GGGCAGATCCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1633
DB 1358 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1411
QY 1634 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1693
DB 1412 GAGGGGTGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1471
QY 1694 GCTTTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1753
DB 1472 GCTTCCAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1531
QY 1754 AGTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1813
DB 1532 AGTCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1591

```

QY	1814	AAGAAACAATATTCCTCTCTGCTGTAAGAGCAACACCTCAGAGCATCAAGACGACATCG	1873
Db	1592	AGGACAGATCTTGTGTGTCATCAAAAGCAAACTCTTAGCTCCATCAAGGCTGACATCG	1651
QY	1874	TCGATGCCCTTTCTGTGCATATCCAGCCAGCCCTGAGTCAACAGTGTCTGTCCACAGACAGCTTCA	1933
Db	1652	TGCACGCTTCTGTGCAATGCCAGTGTCCAGCCAGGTCACAGAGGTCATGCCAAAGAGCTTCC	1711
QY	1934	GGGCGGAGTACAAAGCCAGTGGCGGCCCTCTCTTCCAAAGCCCGCTCCGCTTCCAGG	1993
Db	1712	GGGGCGAGTACAAAGCCAGGCGGGGGCCAGCCGCTGTCCAGAGCCGCTCAAGTTCACAG	1771
QY	1994	TGCACATAGCTCCGTCAGAGGTCOCAGAGCCCTCCCGCCGAGCGGGAGGCGAGCGGAGTG	2053
Db	1772	TTGATATCACTTACACGAGAGGT-----GGGAGGCGCAGAAAGGAGA	1813
QY	2054	GTGGCATTCTACTCCGTCACCTTCACTCTCATCTCGGGGCCAGCCGCTGCTTCAAGCCAG	2113
Db	1814	ACGGCATCTACTCCGTCACCTTCACTCCGTCAGGCCCCAGCCGCTGCTTCAAGAGGG	1873
QY	2114	TGTGTGAGACCATTCAGGACACAGCTCCTGAGACACTCATAGCAGCCCTTCCTGACAGCCC	2173
Db	1874	TGTGTGAGACCATTCAGGACCCAGCTGTGTGAGACACACAGCAGCCCGCTTCGCGCCAGCACT	1933
QY	2174	TGCGACA	2180
Db	1934	TGTCAAGA	1940
RESULT 8			
LOCUS	AX327993	2219 bp	DNA
DEFINITION	Sequence 1 from Patent WO0181588.	Linear	PAT 07-JAN-2002
ACCESSION	AX327993		
VERSION	AX327993.1	GI:18098146	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Meyers, R.		
JOURNAL	2246, protein kinase molecules and uses therefor		
FEATURES	Patent: WO 0181588-A.1 01-NOV-2001;		
source	Millennium Pharmaceuticals, Inc. (US)		
	Location/Qualifiers		
	1..2219		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	53..2077		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD20436.1"		
	/db_xref="GI:18098147"		
	/translation="MTSTGKDGAGAHQAYVGYRLTEKLGQGTGLKLVGHCYTCOKR		
	VAIKIVREKLSIEVLAKVEREIAILKILHPIHLADHYENKRYLYLVLEHVGSE		
	LFQIVYKGGRLTPPEARKEFRQIISALDPCHSHSICHRDLKPEMLLDEKNNTIRADF		
	GMASLQYGDLSLITSCGSPYACPEVIRGKRYGGRKADQWVSCGTYILALVGAALPDD		
	DNLRQLEKRRKRYFHMPHTIPPCOSLWMSSEVDAARRLTLLDHOKHMYIGKNE		
	PEPEQPIPRKVVQIRSLPSELDIDPDVLSHMSLSCFEDRNKLLODLLSEENQRMKTY		
	FLDLRKERYSQEEDLIPRNEIDPRKRVDSPLMRHCKRRBERKSMETLVTDGCGE		
	SPVAPARAIEMAQHGORSISGASGSLSTPLSSPYATPHPSRGSPLPTPKGTPIVH		
	TKPESPGNPENPPSPSYGVGFWARLMSIKNSFTGSPREFHRKLOVTPREMSMLN		
	TPSSPELAKKSMGNFISLEKEROIIVYIKDKPLRSITKADIYHAFUSISLSHSYIS		
	QTSBRATKATCGPAYQKRYKROYDITTYIEGGAQENGTISTYTFLLSGPSHRKRR		
	VVEITIQQLSLTHDPPPAHQHLEPPPAAGLSWAGLKGQAVATVSYESSL"		
BASE COUNT	476 a	716 c	683 g
ORIGIN			344 t
Query Match	41.2%	Score 961;	DB 6; Length 2219;
Matches 1477;	Conservative	0; Mismatches 435;	Indels 195; Gaps 6;

QY	74	CCGACGACGGCCAAATATGTGGGCCCTATCGGCTGGAGAAAGCGTGGCCAAAGGACAGA	133
Db	81	CGCAGCACGGCGACTATGTGGGCGCTTACCGGCTGGAGAAAGCGTGGCCAAAGGACAGA	140
QY	134	CAGGCGTGGTAAACTCTGGGGCTCCACTGCATVCACGGGTTGAGAAAGTCCGCATCAAGATCG	193
Db	141	CAGGCTGTGGTAAAGCTGGGGGTTCACTGGGTACCTGTCCAGAAAGTGGGCATCAAGATCG	200
QY	194	TGAACCGGAGAAAGCTGTGGAAGTCCGATCGGTATGTAAGGTGGAGCGGAGATCGCATCC	253
Db	201	TCAACCCCTGGAAGACTCAGCGAGTGGTGTATGTAAGGTGGAGCGGAGATCGCATCC	260
QY	254	TGAAGCTCATCGAACCACCATGTCTCTCAAGTCTCACAGACGTCTACGAAACAAGAAAT	313
Db	261	TGAAGCTCATTGAGACACCCCGACGTCTTAAAGTGTGACAGACGTTTATGTAACCAAAAAAT	320
QY	314	ATTGTGTACGTGTTCTGGACACAGTCTCGGGGGGTGAGCTATTTGCATCTACGTGTAAAGA	373
Db	321	ATTGTGTACGTGTGTGTAAGAACAGTGTGCAAGTGTGTAGTCTTTGCATCTACGTGTAAAGA	380
QY	374	AGGGGAGACTGACGCCCCAGGAGGCCGAAAGTCTTCCGCCAGATGTGTGTGCGCTGG	433
Db	381	AGGGGAGGCTGACGCCCTTAAGAGGCTCGGAAAGTCTTCCGCCAGATCTCTGCGCTGG	440
QY	434	ACTTCTGCACAGCTACTGTCATGTGCGACAGACCTTAAGGCCCGAAGCCGTGCTTTGG	493
Db	441	ACTTCTGCACAGCACTCCATATGTCACAGGAGATGTGAACCTGTAAACCTCTGCTGG	500
QY	494	ATGAGAAAAACAACATCCGATTTGACAGACTTCCGGATGCGTCCCGCAGGTGGGGGACA	553
Db	501	ACGAGAAAGAACAAATCCTCGCATCGCAGACTTTGGATGGGTGCCCTGCGAGTTTGGGACA	560
QY	554	GCCCTCTGGAGACCAAGTGGGGGTCCCGCCATTATGCGTGTCCAGAGTGTGATTAAGGGG	613
Db	561	GCCGTGTGGAGACCAAGTGTGGGTCCCGCCACTACACCTTCGCCCGAGGTGATCCGGGGG	620
QY	614	AAAATATGTATGGGCGCGCGGCGCAACATGTGGAGCTGTGGAGTACCTCTTCCGCCCTGC	673
Db	621	AGAAAGTATGACGGCCGGGAAGGCGGACGTGTGGAGCTGTGGCGTCACTCTGTTCGCCCTTGC	680
QY	674	TGCGTGGGGGCTCTGCCTTGTATGACAGACAACCTCCGACAGCTGCTGGAGAAAGTGAAC	733
Db	681	TGHTGGGGGCTCTGCCTTGTGACAGATGACAATCTTGGACAGGTGCTGGAGAAAGTGAAC	740
QY	734	GGGGCTGTCCACATGCCCCCACTTATCCCGGCCACTGCAAGATCTGTCTACCGGGGA	800
Db	741	GGGGCTGTGTCCACATGCCCCCACTTATCCCGGCCACTGCAAGATCTGTCTACCGGGGA	800
QY	794	TGATTCGAAGTGGAGCCCGAAAAAAGGCTGAGTGTGAGCAAAATTCAGAAACATTCCTGGT	853
Db	801	TGAGCGAGTGTGAGCGCCGACAGCGCGCTTCAAGTATAGACATTCAGAAACATATGTGT	860
QY	854	ACCTAGCGGGAAAAACAGACGCCAAGACCCGTGTGCTGGAGCCAGGCCCTTGCGCGGGGTAG	913
Db	861	ATATATAGGGGGCAAAAGAACCCCGGAACC-----AGAGACGCCCATTTCTCTCAAGGTGC	914
QY	914	CCATGCGGAAGCTGTCCCATTCACAGGAGAGAGTGGACCCCGACGTCCTTAGAGAGCATGTGCAT	973
Db	915	AGATTCGCTGCTGCCCCAGCTGTGAGGACATTCACCCCGACGTCGTGGACACATGTGCATCT	974
QY	974	CACGTGGGCTGCTTACGAGGACCGGAGAGAGTGCATGCGACGTGCGCACAATGTAGAGAGAGA	1033
Db	975	CACGTGGGCTGCTTCCGAGACCGGACCAACAAGTGTGTGAGAGACGTGCTGTCCGAGAGAGAGA	1034
QY	1034	ACCAAGAAAAGATGATATATTATCTGCTTTTGGATCGAAGAGCGGTATCCAGCTGTG	1093
Db	1035	ACCGAGGAGAGATGATATTACTTCTCTCTCTCGACCGGAAAGAAAGGTACCCGAGCCAGG	1094
QY	1094	AGGACACGAGACCTCCCTCCCGGATAGATGTGTAACCCCGCCGGAAGGTGTGAGATTTCG	1154
Db	1095	AGGATGAGAGACTTCCCGCCCGGACGAGATAGACCTTCCCGGAAGCGTGTGAGACTCC	1154



QY 497 AGAAAAACATCCGATTCGACATTCGCGATTCCTCCAGTGGGGGACAGCC 556  
 Db 241 AGAAGAACATCCGATTCGACATTCGCGATTCCTCCAGTGGGGGACAGCC 300  
 QY 557 TCTGGAGACACATCCGATTCGCGATTCCTCCAGTGGGGGACAGCC 616  
 Db 301 TCTGGAGACACATTCGCGATTCCTCCAGTGGGGGACAGCC 360  
 QY 617 AATATATGAGCCGCGGAGACATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 676  
 Db 361 AGATGACGCGCGGAGAGGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 420  
 QY 677 TGGGGGCTCTGCGCTTCGATGAGACACATTCGCGATTCCTCCAGTGGGGGACAGCC 736  
 Db 421 TGGGGGCTCTGCGCTTCGATGAGACACATTCGCGATTCCTCCAGTGGGGGACAGCC 480  
 QY 737 GCGCTTCACATTCGCGATTCCTCCAGTGGGGGACAGCC 796  
 Db 481 GCGCTTCACATTCGCGATTCCTCCAGTGGGGGACAGCC 540  
 QY 797 TCGAATGAGCCCGGAGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 856  
 Db 541 GCGAGATGAGCCCGGAGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 600  
 QY 857 TTAGCGGAGAGACAGAGCCGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 916  
 Db 601 TTAGCGGAGAGACAGAGCCGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 654  
 QY 917 TCGGAGATTCGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 976  
 Db 655 TCGGAGATTCGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 714  
 QY 977 TGGGCTCTGCGCTTCGATGAGACACATTCGCGATTCCTCCAGTGGGGGACAGCC 1036  
 Db 715 TGGGCTCTGCGCTTCGATGAGACACATTCGCGATTCCTCCAGTGGGGGACAGCC 774  
 QY 1037 AGAAGAACATTCGCGATTCCTCCAGTGGGGGACAGCC 1096  
 Db 775 AGAAGAACATTCGCGATTCCTCCAGTGGGGGACAGCC 834  
 QY 1097 ACCAGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 1156  
 Db 835 ACCAGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 894  
 QY 1157 TCGTGAACCGGAGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 1216  
 Db 895 TCGTGAACCGGAGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 954  
 QY 1217 CGGATTCGCGCTTCGATGAGACACATTCGCGATTCCTCCAGTGGGGGACAGCC 1276  
 Db 955 CGGATTCGCGCTTCGATGAGACACATTCGCGATTCCTCCAGTGGGGGACAGCC 1005  
 QY 1277 ACAGCAGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 1336  
 Db 1006 ACAGCAGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 1065  
 QY 1337 GCAAGCAGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 1396  
 Db 1066 GCAAGCAGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 1078  
 QY 1397 GAGGCGGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 1456  
 Db 1079 GAGGCGGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 1078  
 QY 1457 GAGGCGGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 1516  
 Db 1079 GAGGCGGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 1115  
 QY 1517 CAGGCTCCGCGCTTCGATGAGACACATTCGCGATTCCTCCAGTGGGGGACAGCC 1576  
 Db 1116 CAGGCTCCGCGCTTCGATGAGACACATTCGCGATTCCTCCAGTGGGGGACAGCC 1149  
 QY 1577 CAGGCTCCGCGCTTCGATGAGACACATTCGCGATTCCTCCAGTGGGGGACAGCC 1636

Db 1150 AGAGCCCGGAGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 1203  
 QY 1637 GAGCAGGAGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 1696  
 Db 1204 GAGCAGGAGAGATTCGCGATTCCTCCAGTGGGGGACAGCC 1263  
 QY 1697 TTAGCGGAGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 1756  
 Db 1264 TTAGCGGAGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 1323  
 QY 1757 CTTCCCGGAGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 1816  
 Db 1324 CTTCCCGGAGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 1383  
 QY 1817 AACAAATATTCGCGATTCCTCCAGTGGGGGACAGCC 1876  
 Db 1384 AACAAATATTCGCGATTCCTCCAGTGGGGGACAGCC 1443  
 QY 1877 ATGCTTCGATTCGCGATTCCTCCAGTGGGGGACAGCC 1936  
 Db 1444 ATGCTTCGATTCGCGATTCCTCCAGTGGGGGACAGCC 1503  
 QY 1937 CCGAGTACAAAGGAGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 1996  
 Db 1504 CCGAGTACAAAGGAGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 1563  
 QY 1997 ACATGAGCTTCGCGATTCCTCCAGTGGGGGACAGCC 2056  
 Db 1564 ACATGAGCTTCGCGATTCCTCCAGTGGGGGACAGCC 1605  
 QY 2057 GCATTCACCTTCGCGATTCCTCCAGTGGGGGACAGCC 2116  
 Db 1606 GCATTCACCTTCGCGATTCCTCCAGTGGGGGACAGCC 1665  
 QY 2117 TGAAGACATTCGAGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 2176  
 Db 1666 TGAAGACATTCGAGAGATTCGAGATTCGATTCCTCCAGTGGGGGACAGCC 1725  
 QY 2177 CAGA 2180  
 Db 1726 CAGA 1729

RESULT 10  
 HSM803233  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1014)  
 Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Well, B. and  
 Wiemann, S.  
 Direct Submission  
 Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
 Braunschweig/Germany) within the cDNA sequencing consortium of the  
 German Genome Project.  
 This clone (DKFZ547F069) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
 information about the clone and the sequencing project is available  
 at <http://mips.gsf.de/proj/cDNA/>.  
 Location/Qualifiers  
 1..1014  
 /organism="Homo sapiens"

```

/db_xref="taxon:9606"
/chromosome="19"
/clone="DKFZp547F069"
/tissue_type="brain"
/clone_lib="547 (synonym: hfb1). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="fetal"
1. 1014
/gene="DKFZp547F069"
<1. 541
/gene="DKFZp547F069"
/note="similarity to PEN11B (Homo sapiens)"
/codon_start=2
/product="hypothetical protein"
/protein_id="CAD38595.1"
/db_xref="GI:21732472"
/translation="ISLDKEEJFVLKDKPLISIKADIYAFILSIPLSHVSLSOTS
FRREYKASGSPYFQKRVROVDISSECEPSPRRGSGGGGITYTFTLISGSRK
FRKVFETIQQLSTHDQPSVQALADERNQAQTPAGAPPRILOPPGRPDPELSSSP
RRGPKDKKLALNGTLP"
968
polyA_site
/gene="DKFZp547F069"
229 a 327 c 275 g 183 t
BASE COUNT
ORIGIN
Query Match 23.1%; Score 538; DB 9; Length 1014;
Best Local Similarity 100.0%; Pred. No. 2e-87;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1797 CATCTCTGGACAAAGAAACAATAATTCCTGCTGCTAAAGCAAACTCTGACAG 1856
DB 1 CATCTCTGGACAAAGAAACAATAATTCCTGCTGCTAAAGCAAACTCTGACAG 60
QY 1857 CATCAAGACAGACATCGTCATGCTTCTGTCATGCCCAAGCTGAGTACAGTGTCT 1916
DB 61 CATCAAGACAGACATCGTCATGCTTCTGTCATGCCCAAGCTGAGTACAGTGTCT 120
QY 1917 GTCACAGACAGCTTCAAGGCGGAGTACAGGCAAGTGGGCGCCCTCGTCCAA 1976
DB 121 GTCACAGACAGCTTCAAGGCGGAGTACAGGCAAGTGGGCGCCCTCGTCCAA 180
QY 1977 GCCCTCCGCTTCAGAGTGACATCAGTCTCTGAGAGGTCCAGAGCCCTCCCGGAG 2036
DB 181 GCCCTCCGCTTCAGAGTGACATCAGTCTCTGAGAGGTCCAGAGCCCTCCCGGAG 240
QY 2037 GGAAGGAGGAGAGTGTGATCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2096
DB 241 GGAAGGAGGAGAGTGTGATCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 2097 CGGCGGTTCAAGCGAGTGTGAGACCATCCAGGACAGCTCTGAGCACTCATGACCA 2156
DB 301 CGGCGGTTCAAGCGAGTGTGAGACCATCCAGGACAGCTCTGAGCACTCATGACCA 360
QY 2157 GCCCTCCGCTTCAGAGTGACATCAGTCTCTGAGAGGTCCAGAGCCCTCCCGGAG 2216
DB 361 GCCCTCCGCTTCAGAGTGACATCAGTCTCTGAGAGGTCCAGAGCCCTCCCGGAG 420
QY 2217 CCCACCCGGAAGCTGAGAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2276
DB 421 CCCACCCGGAAGCTGAGAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
QY 2277 CGGCGGAGGCGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2334
DB 481 CGGCGGAGGCGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 538
RESULT 11
AK074411 3156 bp mRNA linear PRI 15-FEB-2002
LOCUS AK074411
DEFINITION Homo sapiens cDNA FLJ3831 fis, clone KAI00425, highly similar to
ACCESION AK074411
VERSION AK074411.1 GI:18677005

```

```

KEYWORDS
SOURCE Homo sapiens ileal mucosa cDNA to mRNA, clone_11b:KAI00425.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Kawakami,T., Noguchi,S., Itoh,T., Shigetani,K., Senba,T.,
Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirose,M.,
Ohmori,Y., Ota,T., Suzuki,Y., Ohashi,M., Nishi,T., Shibahara,T.,
Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3156)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Ohayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction,
5'- & 3'-end one pass sequencing; Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partly supported by Science and Technology Agency).
FEATURES
source
1. 3156
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAI00425"
/tissue_type="ileal mucosa"
/clone_lib="KAI0"
/note="Cloning vector: pME18SFL3"
BASE COUNT 588 a 1075 c 957 g 536 t
ORIGIN
Query Match 21.1%; Score 491.8; DB 9; Length 3156;
Best Local Similarity 61.9%; Pred. No. 4.e-79;
Matches 1075; Conservative 0; Mismatches 372; Indels 290; Gaps 7;
QY 444 CAGCTACTCCATCTGCGACAGACGCTAAAGCCGAGAAACCTGCTTGTGATGAGAAA 503
DB 3 CAGCTACTCCATCTGCGACAGACGCTAAAGCCGAGAAACCTGCTGAGAGAGAA 62
QY 504 CAACATCCGATTTGACAGCTTGGCATGCGCTCCCTGAGTGGGGACAGCTCTTGA 563
DB 63 CAACATCCGATTTGACAGCTTGGCATGCGCTCCCTGAGTGGGGACAGCTCTTGA 122
QY 564 GAGCAGCTGGGGTCCCGCATATATGCGTGTCCAGAGTATTAAGGGGAGAAATATGA 623
DB 123 GAGCAGCTGGGGTCCCGCATATATGCGTGTCCAGAGTATTAAGGGGAGAAATATGA 182
QY 624 TGGCGCCGCGGAGCATGTGAGAGCTGTGAGTATCTCTGCTGCTGCTGCTGCTGCTG 683
DB 183 CGGCGGAGGAGGAGAGTGTGAGAGCTGCGGCTATCTGCTGCTGCTGCTGCTGCTG 242
QY 684 TGTGCCCTTTGATGACGACAACCTCCGCAAGCTGTGAGAGAGTGAAGCGGGCGTCTT 743
DB 243 TGTGCCCTTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 256
QY 744 CCACATGCCCCACTTCAATTCCTCCAGATTGCCAGAGCTCTGAGGAGGAGATCGAAGT 803
DB 257 -----ATGATCGAGGT 267
QY 804 GGAGCCCGAAGAGGCTCAGTCTGAGCAAAATCAGAAACATCTTGTACTAGGCGG 863
DB 268 GGAGCCCGAAGAGGCTCAGTCTGAGCAAAATCAGAAACATCTTGTATAGGCGG 327
QY 864 GAAACAGAGCCAGACCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 923

```

Db 328 CAGAGATGAGCCGACACC-----AGAGCAGCCCACTTCCTCCGAAAGGTCCAGATCCGCTC 381  
 QY 924 CTTCCATTCACAGAGAGAGCTGAGACCCGACGCTCTAGAGACATGGATCACTGGGCTG 983  
 Db 382 GCTGCCAGCCTGGAGGACATGAGACCCGAGAGTGTGGACACATGACATCACTACCTGGGCTG 441  
 QY 984 CTTAGAGGAGCCGAGAGAGCTGATCGGAGCTGGCAGTAGAGAGAGAGACCAAGAGAAA 1043  
 Db 442 CTTCCGAGAGCCGACAAAGAGCTGCTGAGAGTCTGCTCCGAGAGAGAGAGAGAGAGAA 501  
 QY 1044 GATGATATATATATGCTGTTTGGATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1103  
 Db 502 GATGATATATATCTCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 561  
 QY 1104 CTTGCTCCGAG 1163  
 Db 562 CTTGCTCCGAG 621  
 QY 1164 CTTGCTCCGAG 1223  
 Db 622 CTTGCTCCGAG 675  
 QY 1224 CTTGCTCCGAG 1283  
 Db 676 CTTGCTCCGAG 732  
 QY 1284 GAGATCCGAG 1343  
 Db 732 GAGATCCGAG 792  
 QY 1344 AAGAGATCCGAG 1403  
 Db 792 AAGAGATCCGAG 797  
 QY 1404 GAGATCCGAG 1463  
 Db 797 GAGATCCGAG 797  
 QY 1464 CTTGCTCCGAG 1523  
 Db 798 CTTGCTCCGAG 835  
 QY 1524 CTTGCTCCGAG 1583  
 Db 835 CTTGCTCCGAG 876  
 QY 1584 CTTGCTCCGAG 1643  
 Db 876 CTTGCTCCGAG 930  
 QY 1644 CTTGCTCCGAG 1703  
 Db 930 CTTGCTCCGAG 990  
 QY 1704 CTTGCTCCGAG 1763  
 Db 990 CTTGCTCCGAG 1050  
 QY 1764 CTTGCTCCGAG 1823  
 Db 1050 CTTGCTCCGAG 1110  
 QY 1824 CTTGCTCCGAG 1883  
 Db 1110 CTTGCTCCGAG 1170  
 QY 1884 CTTGCTCCGAG 1943  
 Db 1170 CTTGCTCCGAG 1230  
 QY 1944 CTTGCTCCGAG 2003

Db 1231 CAGAGCCAG 1290  
 QY 2004 CTTGCTCCGAG 2063  
 Db 1291 CTTGCTCCGAG 1332  
 QY 2064 CTTGCTCCGAG 2123  
 Db 1332 CTTGCTCCGAG 1392  
 QY 2124 CTTGCTCCGAG 2180  
 Db 1392 CTTGCTCCGAG 1449  
 RESULT 12  
 AB014885 3933 bp mRNA linear INV 06-FEB-1999  
 LOCUS  
 DEFINITION Halocynthia roretzi HtPopK-1 mRNA, complete cds.  
 ACCESSION AB014885  
 VERSION AB014885.1 GI:3172110  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Halocynthia roretzi  
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
 Stolidobranchia; Pyridae; Halocynthia.  
 REFERENCE  
 1 (sites)  
 Sasakura, Y., Ogasawara, M. and Makabe, K.W.  
 Maternally localized RNA encoding a serine/threonine protein kinase  
 in the ascidian, Halocynthia roretzi  
 Mech. Dev. 76 (1-2), 161-163 (1998)  
 JOURNAL  
 MEDLINE  
 98440280  
 2 (bases 1 to 3933)  
 Sasakura, Y., Ogasawara, M. and Makabe, K.W.  
 Direct Submission  
 AUTHOR  
 TITLE  
 JOURNAL  
 FEATURES  
 source  
 1.3933  
 /organism="Halocynthia roretzi"  
 /db\_xref="taxon:7729"  
 /tissue\_type="egg"  
 /dev\_stage="early embryonic"  
 /tissue\_lib="fertilized egg"  
 1.3933  
 /gene="HtPopK-1"  
 166.2262  
 /gene="HtPopK-1"  
 /codon\_start=1  
 /protein\_id="BA28663.1"  
 /db\_xref="GI:3172111"  
 /translation="MSNPQGGVYGVYKLEKTLGKGGTGLVKGVHOMTGKVAVKI  
 VNEKLSSEVYINVKREIATMLKLEHHLGLHLYEKKLYLILEYVSGELEDYL  
 VQKGLPREARREFROIISAVYDCHAHNVCCHRLKLENLIDENKNIKVAIDFGMAST  
 OPGFLFETSCGSHYACPEYIRGERDGNRADVWSCGILFALVGLAPDDDLRO  
 LLEKRVGYHIFHPVPDADONLIRGMDVDPKRLSLQVLOLHPMPGSGNSVGYV  
 VTDPVPVYIDVPLPREESVDPLVLAISLGLGFCNKEKILKILITEONTEKYVYI  
 MLRRKRYSFDDADSLPKHPDAPRKRVDSLSLSSNGDWCVPIDPQRKSAE  
 SLILDTSSPLSLSRKSTETHORQSLTGESNSRLVNCISQITAEKSRINGTPVR  
 GTTCSSNOVPPOINTPASPMPWRQSLSLKNTFMGSPFHRKMQAPSEEDVENGNS  
 SSELKRSWNGENMSRYSSTHEDELPYALAYKNRLLSVKSELVHAFSLIPNLTHS  
 WSLPTRRCDYRSSTGTSVDFHORSIKFOVDIIQHSLSIDQENKRSQDVSFTI  
 ASLISGPIRRYKRYVELLDMQMSIVAAANOQKISGERSAFSHETIOLCKTVNNLHOD  
 NSTWALINPNHKSRYK"  
 3933  
 /gene="HtPopK-1"  
 /note="19 A nucleotides"  
 BASE COUNT 1138 a 803 c 751 g 1241 t  
 ORIGIN



Db 382 TTGCAATTATGAGTGTGATAGACATCCACATGTCCTTATTTGTACGATGCTACGAAA 441  
 QY 305 ACAAGAAATATTTTACTGCTGCTTGGAGACACGCTCTCGGGGGGTGAGCTATTCGATACC 364  
 Db 442 ATTAAGAAATATTTTATTTTGTCTACTAGAAACATGTTAGCGGTGGCGAGCTGTTCGATACC 501  
 QY 365 TGTAAAGAAAGGGGAGCTGACGCGGCAAGAGGCGGCAAGTCTTCCCGCAGATTTGTGT 424  
 Db 502 TTGTGCGGAAAGGACGCGCTGATGTCGAAAGAGCGCGAAAGTCTTCTTCCGTAATCATCT 561  
 QY 425 CTGGGCTGAGCTTCTGCGACAGCTACTCTGCTGCGACAGACCTAAAGCCGAGAACCC 484  
 Db 562 CTGGGCTGAGCTTCTGCGACAGCTACTCTGCTGCGACAGACCTAAAGCCGAGAACCC 621  
 QY 485 TGTCTTTTGGATGAGAAACAAACATCCCATTTGAGACTTCCGATGCGCTCCCTGACAG 544  
 Db 622 TATTATCAGAGAGAGAGATTAACATTAAGATGAGATTTTGGAAATGCTAGCTCTCAAG 681  
 QY 545 TGGGGAGAGAGCTCTGCGAGACAGCTGCGGGTCCCGCATTTATGCTGCGAGAGTGA 604  
 Db 682 TGGAGAGGCTGATTTGGAGACAGCTTGTGATCTCCACATTTATGATGCTCAAGAGTTA 741  
 QY 605 TTAAGGGGAAAAATATGATGAGCGCCGCGGACAGACATGAGCTGAGTATCTCTCT 664  
 Db 742 TTGGGGTGAAGATAGACAGCGTGAAGAAAGAGAGCTTTGAGTTGTGGCTGTATCTCT 801  
 QY 665 TCGCCCTGCTGCTGGGGCTGCTGCCCTTTGATGAGACAACTCCGCCAGCTGCTGAGA 724  
 Db 802 ACGCATTTATTAATGTTGGTGGCTTCCATTTGATGATGAGCAATCTCCGAAATCTTCTGAAA 861  
 QY 725 AGGTGAAGAGGGGCTCTTCCACATGCCCCCATTTCTCCATGCTCCGATGCTCCAGAGCTCC 784  
 Db 862 AAGTGAAGAGGGGCTCTTCCACATGCCCCCATTTCTCCATGCTCCGATGCTCCAGAGCTCC 921  
 QY 785 TGAGGGAATGATGAGAGTGAAGCGCCGAAAAAGGCTCAGCTGAGCAAAATTCAGAAAC 844  
 Db 922 TGGCGGCAATGATGAGAGTGAAGCGCGGAAAAAGATATTCCTAGCCGAGCTGTCAAAAC 981  
 QY 845 ATCTGTTGACTGCTGAGGGGAAACAGAGCCGAGACCTGCTGAGAGCCGCTGAGCC 904  
 Db 982 ATCCATG--GTATCCGGAACAACAAAGGAGCAATCTGAAATGAGCTTCCATGCTCAC 1038  
 QY 905 GCGGGTATGAGCAGCTGAGAGCTGCTCCATCCAGAGAGCTGAGAGCCGAGCTCAGAGA 964  
 Db 1039 AAGTTGAGCAGAGCAGCTATACCTGCTGAGAGATTCATGCTGCTGCTGCTGAGCC 1098  
 QY 965 GCATGCAATCAGCTGCTGCTGCTGAGAGCCGAGAGCTGAGAGCTGAGAGCTGAGAG 1024  
 Db 1099 ATATGAACTGCTGAGATGTTTCAAGACAAACAGAACTGATTAATGAATATTGTGCG 1158  
 QY 1025 AGGAGGAGAACCAAGAAAGATGATATTTATCTGCTTTTGGATGCGAAGAGCGGTATC 1084  
 Db 1159 CAAACACAACTGAAAGAGTGTATCTTCTCTCTTCTGCTGAAAGAGAGCGGCGCC 1218  
 QY 1085 CCAGCTGTGAGAGCA 1100  
 Db 1219 CTGACAGAGAGAGCA 1234

RESULT 14  
 LOCUS AY060288 5609 bp mRNA linear INV 05-NOV-2001  
 DEFINITION Drosophila melanogaster GH13047 full length cDNA.  
 ACCESSION AY060288  
 VERSION AY060288.1 GI:16648123  
 KEYWORDS  
 SOURCE Drosophila melanogaster.  
 ORGANISM Drosophila melanogaster.  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Ephydroidea; Drosophila.  
 REFERENCE 1 (bases 1 to 5609)  
 AUTHORS Stapleton, M., Broksstein, P., Hong, L., Agbayani, A., Carlson, J.,

TITLE  
 JOURNAL  
 COMMENT  
 Submitted (29-OCT-2001) Berkeley Drosophila Genome Project,  
 Lawrence Berkeley National Laboratory, One Cyclotron Road,  
 Berkeley, CA 94720, USA  
 Sequence submitted by:  
 Lawrence Berkeley National Laboratory

FEATURES  
 source  
 This clone was sequenced as part of a high-throughput process to  
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,  
 Science 2000). The sequence has been subjected to integrity checks  
 for sequence accuracy, presence of a polyA tail and contiguity  
 within 100 kb in the genome. Thus we believe the sequence to  
 reflect accurately this particular cDNA clone. However, there are  
 artifacts associated with the generation of cDNA clones that may  
 have not been detected in our initial analyses such as internal  
 priming, priming from contaminating genomic DNA, retained introns  
 due to reverse transcription of unspliced precursor RNAs, and  
 reverse transcriptase errors that result in single base changes.  
 For further information about this sequence, including its location  
 and relationship to other sequences, please visit our web site  
 (http://fruitfly.berkeley.edu) or send email to  
 cdna@fruitfly.berkeley.edu.  
 location/Qualifiers  
 1..5609  
 /organism="Drosophila melanogaster"  
 /strain="y; cn bw sp"  
 /db\_xref="taxon:7227"  
 /map="72A5-72B1"  
 1..5609  
 /gene="CG6114"  
 /note="Alignment with genomic scaffold AE003529"  
 /db\_xref="FLYBASE:FBgn0036544"  
 272..2377  
 /gene="CG6114"  
 /note="Longest ORF"  
 /codon\_start=1  
 /product="GH13047P"  
 /protein\_id="AA25327.1"  
 /db\_xref="GI:16648124"  
 /translation="MASIOPAGSMLEFSCSPHYACPEVINGEKYDGRKADVMSGVI  
 LVALLVAGLPEDDNLRLLEKVRGPHVPHVPPDCOSILRMIVNPPRLTAE  
 INRPVWTAGGKELELEPMYVYOHVHPTAANDPVYATACSGCFREKELIO  
 ELLSSHNTEKVIYFLLERKRRPALEDEDEIAOKRSELDADVPKRLDTRKING  
 TNAESTGQISEGSLTPRQAFNRSTSTRNHRSPPTVTSVRSSTSSPTKNS  
 PMSAQOQANALISRPSPAGTRSHYGDNRGSHSVSTPSHSOKSIEGSPVTV  
 REPRIRRDGSLROERGSGSPDRDGCIPGSPGNSGSGTASPVYHNRNGLPTIA  
 ISMFHPDPSNSVYVNPSPMNNSSPCMPSPCTPGGOLMKTLTNRKNFGLSPR  
 HRRKQVSADEVILTPRESSPELTKRSFNGFLITTEKETETIYIKGRPIAYKHLIH  
 APLSMALSHSVYSPFSFRYRYKRNKGFPYRQRYKQYDIAICOGDADLAFAL  
 TTLLSGNTRRRICEHIOYVCSKRPDPSPPTVSTVQAVSESSSCGVSSERL  
 SYKROYIENDMENDSLFYSKSGNGRSTNNNNANVDIPSPIAVRSNSETAHER  
 NRELOSERQATMTWSGAIA"

CDS  
 gene  
 BASE COUNT 1803 a 1224 c 1164 g 1418 t  
 ORIGIN

Query Match 16.0%; Score 373.6; DB 3; Length 5609;  
 Best Local Similarity 65.5%; Pred. No. 9.2e-58;  
 Matches 563; Conservative 0; Mismatches 294; Indels 3; Gaps 1;  
 QY 259 CTCATGACACACCACATGCTCCAGCTCCAGAGCTACGAGAAAGAAATATTG 318  
 Db 2 CTAATGATATATCCACACGCTCTGCGCTGAGCATGTGTGAGAAAGAAAGATTG 61  
 QY 319 TACCTGTTCTGAGCAGCTCTGCGGGGTGAGCTATTCGACTACCTGTAAGAGGGG 378  
 Db 62 TATTGATATTGAGCATGTATCCGCGAGAGAGCTCTCGATTACCTGTGAAGAGGT 121



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 12:29:03 ; Search time 19 Seconds  
(without alignments)  
1204.791 Million cell updates/sec

Title: US-10-003-690-2

Perfect score: 4108  
Sequence: 1 MSGAKEGGGSPAYHLPH.....PRNGPPKRLTLATNGPLP 778

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgnt2\_6/prodata/1/1aa/5A.COMB.pep:\*
- 2: /cgnt2\_6/prodata/1/1aa/5B.COMB.pep:\*
- 3: /cgnt2\_6/prodata/1/1aa/6A.COMB.pep:\*
- 4: /cgnt2\_6/prodata/1/1aa/6B.COMB.pep:\*
- 5: /cgnt2\_6/prodata/1/1aa/PCMTUS.COMB.pep:\*
- 6: /cgnt2\_6/prodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2799	68.1	668	4	US-09-930-181-2
2	2490.5	60.6	603	4	US-09-930-181-17
3	2385.5	58.1	595	4	US-09-930-181-4
4	827	20.1	745	4	US-09-523-849-36
5	795	19.4	729	2	US-08-677-298-2
6	795	19.4	729	4	US-09-523-849-33
7	788	19.2	733	4	US-09-523-849-32
8	784	18.1	776	4	US-09-523-849-34
9	753.5	18.3	633	4	US-08-557-006C-43
10	713.5	17.4	345	3	US-09-101-146-1
11	715	17.4	552	4	US-08-557-006C-40
12	700.5	17.1	257	3	US-09-101-146-6
13	695	16.9	604	4	US-09-523-849-35
14	691	16.8	257	2	US-07-857-224B-25
15	570	14.1	149	4	US-09-930-181-18
16	543.5	13.2	222	2	US-07-857-224B-26
17	522.5	12.7	436	4	US-09-734-673-2
18	521	12.7	436	4	US-09-523-849-2
19	521	12.7	1050	4	US-09-428-711A-16
20	518.5	12.6	353	3	US-08-688-988-31
21	512	12.5	290	4	US-09-734-673-4
22	512	12.5	334	4	US-09-523-849-31
23	510	12.4	1037	4	US-09-428-711A-21
24	509.5	12.4	260	2	US-07-857-224B-27
25	495	12.0	339	3	US-08-688-988-33
26	493	12.0	351	3	US-08-688-988-28
27	491	12.0	354	3	US-08-688-988-29

28	490.5	11.9	260	2	US-07-857-224B-28	Sequence 28, Appl
29	482.5	11.7	370	2	US-08-878-989-19	Sequence 19, Appl
30	482.5	11.7	370	4	US-09-272-796-19	Sequence 19, Appl
31	482.5	11.7	370	4	US-09-457-040B-31	Sequence 31, Appl
32	482.5	11.7	1051	4	US-09-428-711A-14	Sequence 14, Appl
33	482	11.7	556	4	US-09-800-960-4	Sequence 4, Appl
34	477.5	11.6	264	2	US-07-857-224B-18	Sequence 18, Appl
35	477.5	11.6	339	3	US-08-688-988-32	Sequence 32, Appl
36	477.5	11.6	363	3	US-08-688-988-30	Sequence 30, Appl
37	475.5	11.6	565	4	US-09-800-960-2	Sequence 19, Appl
38	474.5	11.6	264	2	US-07-857-224B-19	Sequence 23, Appl
39	469.5	11.4	295	1	US-07-951-715A-23	Sequence 23, Appl
40	469.5	11.4	295	2	US-08-459-448A-23	Sequence 23, Appl
41	469.5	11.4	295	3	US-08-459-595A-23	Sequence 23, Appl
42	469.5	11.4	295	3	US-08-459-504B-23	Sequence 23, Appl
43	469.5	11.4	295	3	US-08-459-444-23	Sequence 23, Appl
44	469.5	11.4	295	4	US-09-547-422-23	Sequence 23, Appl
45	467.5	11.4	339	3	US-08-688-988-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-09-930-181-2  
Sequence 2, Application US/09930181  
Patent No. 6455292  
GENERAL INFORMATION:  
APPLICANT: Origene Technologies  
TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas  
FILE REFERENCE: 160 101 VI  
CURRENT APPLICATION NUMBER: US/09/930,181  
CURRENT FILING DATE: 2001-08-16  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 668  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-930-181-2

Query Match 68.1%; Score 2799; DB 4; Length 668;  
Best Local Similarity 75.1%; Pred. No. 8.9e-171;  
Matches 348; Conservative 54; Mismatches 48; Indels 80; Gaps 8;

QY	1	MSGAKEGGGSPAYHLPHPHPOHAQYVPRLETKKGOTGLVGHCTGQKV	60
DB	1	MTSTGADGA-----QHAQYGPRLRETKKGOTGLVGHCTGQKV	45
QY	61	AIKIVNREKLSSEVLMKVEREIALIKLIEHPVILKLDHYENKKYLVLEHVSGLLED	120
DB	46	AIKIVNREKLSSEVLMKVEREIALIKLIEHPVILKLDHYENKKYLVLEHVSGLLED	105
QY	121	YLVRKGLTPREKARKFRQVLSALDFCHSYISICRDLKPEMLLDKNNIRIADFGASL	180
DB	106	YLVRKGLTPREKARKFRQVLSALDFCHSYISICRDLKPEMLLDKNNIRIADFGASL	165
QY	181	QVDSILLETSCGSPHYACPEYIKGEKDGRRADMMSCVILFALLVGLPPDDDLRLQL	240
DB	166	QVDSILLETSCGSPHYACPEYIKGEKDGRRADMMSCVILFALLVGLPPDDDLRLQL	225
QY	241	EKKVKGVFHMPHFLPPDCQSLRGMIEVEPERKLSLEQIQHPWYLGKHEPDPLEPAP	300
DB	226	EKKVKGVFHMPHFLPPDCQSLRGMIEVEPERKLSLEQIQHPWYLGKHEPDPLEPAP	284
QY	301	GRVYAMSLPSNGELDDVLESASLGCFRREKLHRLSEENQEMITVLLDRKER	360
DB	285	RRVQIHSPLSLIEDIDVDLSMISLGCFRDRNRLKLDLSEENQEMITVLLDRKER	343
QY	361	YPCSEDDLPDRNDVDPKRRKVDSPMLSRHGRPERKSMELSTTDGSGGSPVPTARA	420
DB	344	YPCSEDDLPDRNDVDPKRRKVDSPMLSRHGRPERKSMELSTTDGSGGSPVPTARA	400

QY 421 LEMAHORSRSVSGASTGLSSPLSSPRPVFSFSPERGADGAGGSPSTKTQTLPS 480  
DB 401 LEMAHGORSRSISGASSGLSTSPSSPR----- 429  
QY 481 RGRGCGAGCQPPPSARSPLPGPPSPSSGGGTLPLSHPLTPRASPTGTPPPSP 540  
DB 430 -----VTPHPSRGSPLPTPKG-----TPVHTPKESAGTNPPTPPSP 468  
QY 541 GCGVGGAAMRSRLNSIRNSFLSGSPRRRRKMQVPTAEEMSSLTPESSPELAKRSWFGNTI 600  
DB 469 --SVGVPMFARLNSIKNSFLSGSPRRRRKMQVPTAEEMSSLTPESSPELAKRSWFGNTI 526  
QY 601 SLDKESQIFLYLKDPLSSIKADIVHAFPLSTPSLSHVSLSGTSFRAEKAKAGSPVFOKP 660  
DB 527 SLEKEQIFLYLKDPLSSIKADIVHAFPLSTPSLSHVSLSGTSFRAEKAKAGSPVFOKP 586  
QY 661 VRFQVDSISCEPSPRRRSGGGGCIYSVFTLISGSPRRRRKRVETIOAQLSTHOP 720  
DB 587 VKQVDTITTEGGEAOK-----NGIYSVFTLISGSPRRRRKRVETIOAQLSTHOP 640  
QY 721 SVQALDEKN 730  
DB 641 AAOHLSDTTP 650

## RESULT 2

US-09-930-181-17  
; Sequence 17, Application US/09930181  
; Patent No. 6455292  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies  
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas  
; FILE REFERENCE: 16U 101 VI  
; CURRENT APPLICATION NUMBER: US/09/930,181  
; CURRENT FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 17  
; LENGTH: 603  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-930-181-17

Query Match 60.6%; Score 2490.5; DB 4; Length 603;  
Best Local Similarity 72.8%; Pred. No. 3.6e-151;  
Matches 488; Conservative 52; Mismatches 45; Indels 85; Gaps 9;

QY 87 LLEHPRVLTLDHYENKKYLYLVLEHVSGLLPDYLVKKGRITPKARKFROIVSALDF 146  
DB 1 LLEHPRVLTLDHYENKKYLYLVLEHVSGLLPDYLVKKGRITPKARKFROIVSALDF 60  
QY 147 CHSYSTICHRDLKPELNLDEKNNIRIADFGMASLQVDSILETSCGSPHYACPEVIKGEK 206  
DB 61 CHSHSTICHRDLKPELNLDEKNNIRIADFGMASLQVDSILETSCGSPHYACPEVIKGEK 120  
QY 207 YGRRADAMSCGYITLALVGAALPDDDNIRQLLEKVKGVFHPHFIPDDCSILLRMTI 266  
DB 121 YGRRADAMSCGYITLALVGAALPDDDNIRQLLEKVKGVFHPHFIPDDCSILLRMTI 180  
QY 267 EYEPERLSTLEQIOKHPWYLGKHEPDCLEPARGVAMRSLSNGLDDVLESNASL 326  
DB 181 EYDAARLTLEHIOKHIWYIGKNEPEP-RQPIP-RKVQIRSLPSLEDDIDVDLSMHSI 238  
QY 327 GCFRDERLTARELSEENOEKMIYLLDRKERYPSCEDODLPRRNDVDPKRKRVDSPM 386  
DB 239 GCFRDERLTARELSEENOEKMIYLLDRKERYPSCEDODLPRRNDVDPKRKRVDSPM 298  
QY 387 LSRHGRARRERSMEVLSITDAGGSPVPTRRALMAHORSRSVSGASTGLSSPLS 446  
DB 299 LNRHGRARRERSMEVLSITD--GGSPVPARAIEAHGORSRSISGASSGLSTSPLS 355  
QY 447 SPRSVPFSFSPERGADGAGGSPSTKTQTLPSRGRGAGGAGCQPPPSAKSTPLPGPP 506  
DB 447 SPRSVPFSFSPERGADGAGGSPSTKTQTLPSRGRGAGGAGCQPPPSAKSTPLPGPP 506

DB 356 SPR-----VTPHPSRGSPLPTPK 374  
QY 507 GSPRSGLGTLPLSHPLTPRASPTGTPPPSPGGVGGAAMRSRLNSIRNSFLSGSPRF 566  
DB 375 G-----TPVHDPKSPAGTNPPTPPSP--SVGVPMFARLNSIKNSFLSGSPRF 421  
QY 567 HRRKQVPTAEEMSSLTPESSPELAKRSWFGNTISLDKEQIFLYLKDPLSSIKADIVH 626  
DB 422 HRRKQVPTAEEMSSLTPESSPELAKRSWFGNTISLDKEQIFLYLKDPLSSIKADIVH 481  
QY 627 AFLSTPSLSHVSLSGTSFRAEKAKAGSPVFOKPVRFQVDSISCEPSPRRRSGGGG 686  
DB 482 AFLSTPSLSHVSLSGTSFRAEKAKAGSPVFOKPVRFQVDSISCEPSPRRRSGGGG 535  
QY 687 IYSVFTLISGSPRRRRKRVETIOAQLSTHOPSVQALDEKNCAQTPRAGAPRSLQ 746  
DB 536 IYSVFTLISGSPRRRRKRVETIOAQLSTHOPPAQHLS-----EP 577  
QY 747 PPRGPPPLS 756  
DB 578 PP--PAPGLS 585

## RESULT 3

US-09-930-181-4  
; Sequence 4, Application US/09930181  
; Patent No. 6455292  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies  
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas  
; FILE REFERENCE: 16U 101 VI  
; CURRENT APPLICATION NUMBER: US/09/930,181  
; CURRENT FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-930-181-4

Query Match 58.1%; Score 2385.5; DB 4; Length 585;  
Best Local Similarity 74.4%; Pred. No. 1.7e-144;  
Matches 464; Conservative 51; Mismatches 44; Indels 65; Gaps 7;

QY 107 YVLEHVSGLLPDYLVKKGRITPKARKFROIVSALDFCHSYSTICHRDLKPELNLDE 166  
DB 9 YVLEHVSGLLPDYLVKKGRITPKARKFROIVSALDFCHSYSTICHRDLKPELNLDE 68  
QY 167 KNNIRIADFGMASLQVDSILETSCGSPHYACPEVIKGEKYDGRADAMSCGYITLALV 226  
DB 69 KNNIRIADFGMASLQVDSILETSCGSPHYACPEVIKGEKYDGRADAMSCGYITLALV 128  
QY 227 GALLPDDDNIRQLLEKVKGVFHPHFIPDDCSILLRMTIEVEPERLSTLEQIOKHPWY 286  
DB 129 GALLPDDDNIRQLLEKVKGVFHPHFIPDDCSILLRMTIEVEPERLSTLEQIOKHIWY 188  
QY 287 GSKHEPDCLEPARGVAMRSLSNGLDDVLESNASLGCGRDRRLARELSEENQ 346  
DB 189 GSKNEPEP-EQPIP-RKVQIRSLPSLEDDIDVDLSMHSIGCGRDRRLARELSEENQ 246  
QY 347 EKMITYLLDRKERYPSCEDODLPRRNDVDPKRKRVDSPMISRHGRARRERSMEVLSIT 406  
DB 247 EKMITYLLDRKERYPSCEDODLPRRNDVDPKRKRVDSPMISRHGRARRERSMEVLSIT 306  
QY 407 DAGGGSPPVPTRRALMAHORSRSVSGASTGLSSPLSSPPSPVFSFSPERGADGAG 466  
DB 307 D--GGSPVPARAIEAHGORSRSISGASSGLSTSPSSPR----- 346  
QY 467 GGSPTSKTQTLPSRGRGAGGAGCQPPPSAKSTPLPGPPGSPRSGLTPLSHPLTPRA 526  
DB 347 -----VTPHPSRGSPLPTPKG-----TPVHTPK 371

```

OY 527 SPTGPGTTPPPSCGGGAGMARSLNRSIRSFSGSPFHRKKMOVPFAEEMSSLTIPS 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 372 SPACTPNTPTPPSSP--SYGVFWPRRLNLSIKSISFGSPFHRKKLOVPPEEMSSLTIPS 429
OY 587 SPELAKRSMPFNFTSLDDEQIFVLAKDKPLSSIKADIYHAFSLSPSLSHSVLSQTSFPA 646
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 SPELAKRSMPFNFTSLDDEQIFVLAKDKPLSSIKADIYHAFSLSPSLSHSVLSQTSFPA 489
OY 647 EYKASGGSVYQKPYRFOVDYISSSGPEPSPRRDGGGGGCIYSVTFTLISGSPRRFRVY 706
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 490 EYKATGCGAVYQKPYRFOVDYITVTEGGEAOKR-----NGIYSVTFTLISGSPRRFRVY 543
OY 707 ETIOALLSTHDOPSYVALADEKN 730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 544 ETIOALLSTHDOPSYVALADEKN 567

RESULT 4
US-09-523-849-36
; Sequence 36, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scacchi, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523, 849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: GenBank Accession No. 6458561 g1749794
US-09-523-849-36

Query Match 20.1%; Score 827; DB 4; Length 745;
Best Local Similarity 30.9%; Pred. No. 4.5e-45;
Matches 245; Conservative 114; Mismatches 259; Indels 176; Gaps 26

OY 30 YVGPRLKRTKGGTGTGLVKGVCIGTGGKAAIKTYNREKLSSESVLMKYERETAILKTE 89
    ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16 HIGNRLKLTGKGKGFAPVLAHRLITGGEVAVKTIIDKQLSSLOKLFREVRIKVLN 75
OY 90 HPVYLKLDVYENKKYLLVLEHVSGGELFDVLYMKGRLTPEARKREFFOIYSALDFCHS 149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 HPNVLKLEVELETETLLVMEYASGGEYFDVLAHGRKKEKARKKFOIYSAVOYCHQ 135
OY 150 YSICHRDLKPEMLDKEFNNTIRIADFGMASLOVGSLSLETSCGSPHYACPEVYKGEKYG 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 KEIYARLDKAERLLDLADAMNIRIADFGMSNETFGKIDTFCGSPHYAPYAELEFGKKYGG 195
OY 210 RADWMSGCVILFALLVGLPPDDNLRQLEKVRGVFHPHFIPDCCOSLIRGMEIYE 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 PEVDWMSGLVILYTLVSGSLPDCGNLKEKRLBRVLRGKYRIPEYMSSTDENLTKFELIN 255
OY 270 PEKRLSLIOIKHPYLGKKEHPD---PCLTEAPGRVYMRSLPSNGELDPPVLESMAST 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 PSKRSTLQIMKDRMANYG--HEDELKPYEVLPOTK-----DPRRTELMVSN 302
OY 337 GCFRDRERLHRELSEENOEKMIYVLLLDKREKRPSCEDODLPPRNDYV----- 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 G--YRRELIQSLVGGQRN-EVMATYLLLGKYSLELGGTITLLKPPRADLTNNSAQFPS 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 377 -----PPKRAYDS-----PMLSHKG-----RRPE-----RKSMM----- 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db      360 HKVQSRVSANFKQRRFSDQAGPALPTNSISYKTKQSNNAENKRPEDRESGRAASSTAKY 419
OY      401 -----EVLSTIDAGGGGSPVPTRRALENA--OHQSORSRVSQ--A 436
Db      420 PASPLPGLERKKTPTPTSTNSVLTSTNRNSNPSLLERASLGQASIONGDSLMPESRA 479
OY      437 STGLSSPSLSPR-----SPVSEFSPBPAGDEARGGSGFTSKTQTLPSRGP----- 483
Db      480 STASASAAVSANRAPHROHOKSMSAVHNNKASGLPTESNCEVPRPSTAPRVPVAPSPAH 539
OY      484 ---KGGAGGQ---PPPSASTPLRPPGSPRSGGTPLHSLPLHTRASPTGTPPTP 537
Db      540 NISSGGAPDNTNFPBGVSSSTIHAQRLROYRQOQNLPy-----GVT-P 563
OY      538 PSPGGGVGAGMARSRLNIRNSPLFGSPFHRKKQVPTAAE--MSLTP----- 584
Db      564 ASPSGHSGR--KQASGSIFSKF--TKSFYRNLNPEESKDRTVELTPRHYVSGGDNKEK 639
OY      585 ---ESSPELAKRSM--FGNFTSLDKEROQIFVLKDKPLSSIKADIVHAFTSLPSLSHYL
Db      640 EEFREAKPRSLRFTWMSKTTSSMBPENNEMREIRKVLVLANSOQSLEHBKMYML--CMHGTP 697
OY      640 SOTSRAETAKASGGPSVQKRPYQOVDISSEGEPEPBRDGSGGGTYVTFLISGPS 699
Db      698 GHEDF-----VQEMEEVCK-----LPRSLNG-----VREFRISGTS 729
OY      700 RFRKVVETIOAOL 713
Db      730 MAFKNIAKINTEL 743

```

```

US-08-677-298-2
: Sequence 2, Application US/08677298
: Patent No. 5863729
:
: GENERAL INFORMATION:
: APPLICANT: Pwmlca-Worms, Helen
: TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN TCaK-1
: TITLE OF INVENTION: KINASE
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
: STREET: 5370 Manhattan Circle, Suite 201
: CITY: Boulder
: STATE: CO
: COUNTRY: USA
: ZIP: 80303
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA
: APPLICATION NUMBER: US/08/677, 298
: FILING DATE: 09-JUL-1996
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Caruthers, Jennie M.
: REGISTRATION NUMBER: 34,464
: REFERENCE/DOCKET NUMBER: 9-96
: TELECOMMUNICATION INFORMATION
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 729 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
US-08-677-298-2

```





```

Db 513 FSASEGAGAGGATGATGCTGSSFPVLASPLGS-----OSAPVLOSQGLGATVLP 566
OY 589 -----ELAKRSWGFNFISLDKEQIPLVLDKPLSSIKADI 624
Db 567 PVSPQEGRRASDTSLTGGLKAFRQQLRKNARTKGFGLNK----- 606
OY 625 VHAFLSLPSLSHVSLSQTSFPAEKASGSPVQKPVPRQVDISSEGEPEPRDGS 684
Db 607 -----TKGLARQV-COSSING-----SRGGMSTFHTPA-----PSSGLG--- 640
OY 685 GGIVSVFTLLSGSRPRKRVETI--OAOILSTHDQSVQALADEKNGAOTRPAAGAPR 742
Db 641 -----CTASSRGRSLLEVLVHQRLQLQAHSAVS--SDYQAPQLSPV---PY 685
OY 743 SLOPPPG 749
Db 686 VLFPCDG 692

```

## RESULT 9

```

US-08-557-006C-43
; Sequence 43, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
; APPLICANT: Bertl, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forde, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAF/PHM37588/US
; CURRENT APPLICATION NUMBER: US/08/557,006C
; CURRENT FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 43
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Yeast
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(633)
; OTHER INFORMATION: Yeast SNF1 polypeptide
US-08-557-006C-43

```

```

Query Match 18.3%; Score 753.5; DB 4; Length 633;
Best local similarity 32.8%; Pred. No. 1.8e-40;
Matches 166; Conservative 103; Mismatches 144; Indels 93; Gaps 14;

```

```

OY 14 AYHLPRHPHPPROHAQ-----YVGPYREKLTGKQOTGLYKIGVHCI 55
Db 17 SHHHHHHHHHHHHGGGNSLTNNPKSSLDGAHIGYQIVTKTLEGGSPGVKLAUHTT 76
OY 56 TGOQVAIKIYNREKLSVLM-KYERETAILKLEHPVLKLDHYENKKYUYLEHVS 114
Db 77 TGOQVAKIKIKVLAQSDMGRIERETSYRLRLRHPIIKLYDKSKDELIMVIEY-A 135
OY 115 GGELFDLVKGRITLPREARKFRQIVSALDFCHSYSGICHDLPENLLDEKNNIRIAD 174
Db 136 GNEFLFDIYVRDKMSEQEARFEOIISAVEYCHRHKIVHDLKPEMLLDLHNVKIAD 195
OY 175 FGMAASLOVGSILETSCGSPHYACPEVIKGEKYGRADAMSCGYTLALLVGAIPDD 234
Db 196 FGLSNITDGNFNTKSCGSPHYAPEVYISGLYAGPEVDVMSCGVILYMLCRRLPDE 255
OY 235 NLKQILEKVRGVFHMHFIPPCDSSLGMIIEVEPEKRLSLEQIOKHPWYLGKHE--- 291
Db 256 SIYVLPKRNISNGVYTLTKFLSPGAAGLIKRLILVPLNRISIHIMODMFKVLDPEYIL 315

```

```

OY 292 -PDCLEPAPGRRYAMR-----SLPSNGELDPVLESMAISGCFRDRERLRLRSEEE 344
Db 316 PPD--LKPHDEEENENDSKDGGSPDNDEIDMLVNLISSTMGY-EKDEIYESLESSED 372
OY 345 -----NOKMTIYLLLDKREKRPSCEDODLPPRRNV-----DPP-----RK 380
Db 373 TPAPNEIRDAIYMLIKENKSLI-----KDMKANKSVSDLEDFLSQSPPTFOQSKHQS 427
OY 381 RVDSFMLSRRGKRPRRPERKSMSEVLSTIDAGGSGSPVPTFRALF-----MAQHSQRSVSGA 436
Db 428 QVHEFAKQARNR-----ASAKITQARTYHQSFPMDQYKNEEDSTVSL 470
OY 437 STGL-----SSSPLSPNSPV 452
Db 471 PTSLPQIRANMLAOGSPASAKISPL 496

```

## RESULT 10

```

US-09-101-146-1
; Sequence 1, Application US/09101146
; Patent No. 6124125
; GENERAL INFORMATION:
; APPLICANT: Dartmouth College, St. Vincent's Institute of
; APPLICANT: Medical Research, Kemp et al.
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 66 E. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PC
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/101,146
; FILING DATE: October 7, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PN7450
; FILING DATE: 8 JAN 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: DC-0050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (856) 810-1515
; TELEFAX: (856) 810-1454
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345
; TYPE: Amino acid
; TOPOLOGY: linear
US-09-101-146-1

```

```

Query Match 17.4%; Score 715.5; DB 3; Length 345;
Best local similarity 40.7%; Pred. No. 2.3e-38;
Matches 146; Conservative 65; Mismatches 105; Indels 43; Gaps 5;

```

```

OY 31 VGPYREKLTGKQOTGLYKIGVHCITGQVAKIVNREKLSSEVLMYEREITATIKLEH 90
Db 12 IGHVILIGDTLGVGTFGVKVGKHELTGKVAVKTLNRQKRLDVGKIRREIQMLKFRH 71
OY 91 PHLKLDHYENKKYUYLEHVS GGELFDLVKGRITLPREARKFRQIVSALDFCHSY 150
Db 72 PHIKIKLYQVISTSDIFNMVAYVSGELFDYICNGRLDKESRRLDQQLISGVYDCHRH 131
OY 151 SICHRDLKPEMLLLDEKNNIRIADFGMAASLOVGDLSLETSCGSPHYACPEVIKGEKYDGR 210

```

Db 132 MVAHRLDKPENVLLDAMNNAKIDAFGLSNMMSDGEFLRTSCGSPRYAPAEYISGLYAGP 191  
Qy 211 RADMMSCGVLFFALLVGLALPPDDNLRQLLEKVRGVFHMFFIPDCOSLLRGMEYEP 270  
Db 192 EVDIMSSGVLIALLCGLTFPDDHVPFLFKKIDGIFTYQYINPVSISLKHMLQYDP 251  
Qy 271 EKRLSLEQIOKHPWYLGKHEPDPCLPAPGRVAMRSLSNGELDPVLESMAISLGCGR 330  
Db 252 MKRATIKDIREHEWF---KQDLPSYLFPE-----DPSYSSTMI----- 286  
Qy 331 DRELRHRELSEENOEKMI-----YYLLDRKERYPCGEODL---PP 371  
Db 287 DDEALKEVCEKFCSESEVLSCLYNRHQDPLAVAYHLIDNRIRIMNAKDFYATSP 345

RESULT 11  
US-08-557-006C-40  
; Sequence 40, Application US/08557006C  
; Patent No. 6238547  
; GENERAL INFORMATION:  
; APPLICANT: Berl, Rajindar K.  
; APPLICANT: Carling, David  
; APPLICANT: Forde, Robert A.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
; FILE REFERENCE: NGAP/PM37588/05T  
; CURRENT APPLICATION NUMBER: US/08/557, 006C  
; PRIOR FILING DATE: 1996-03-06  
; PRIOR APPLICATION NUMBER: PCT/GB94/01093  
; PRIOR FILING DATE: 1994-05-20  
; PRIOR APPLICATION NUMBER: GB 9310489.1  
; PRIOR FILING DATE: 1993-05-21  
; PRIOR APPLICATION NUMBER: GB 9318010.7  
; PRIOR FILING DATE: 1993-08-31  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 40  
; LENGTH: 552  
; TYPE: PRT  
; ORGANISM: Rat  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (1)..(1747)  
; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -  
; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with  
; OTHER INFORMATION: nucleotide 1765  
US-08-557-006C-40

Query Match 17.4%; Score 715; DB 4; Length 552;  
Best Local Similarity 40.7%; Pred. No. 4.3e-36;  
Matches 150; Conservative 71; Mismatches 134; Indels 14; Gaps 5;

Qy 31 VGYRLKLTGKGTGGLVLCVHCITGQKVAIKIYVREKL-SESYLKVREITAIKLE 89  
Db 13 IGVYLVGDTLVGTFGVKKGHEHDTGKVAIKIYVREKL-SESYLKVREITAIKLE 89  
Qy 90 HPHVYLVGDTLVGTFGVKKGHEHDTGKVAIKIYVREKL-SESYLKVREITAIKLE 149  
Db 73 HPHVYLVGDTLVGTFGVKKGHEHDTGKVAIKIYVREKL-SESYLKVREITAIKLE 132  
Qy 150 YSICHLRLKPNLILDEKNNIRIDAFGNASLOVGSLSLETSCGSPHYACPEYISGLYAGP 209  
Db 133 HMYVHRLDKPENVLLDAMNNAKIDAFGLSNMMSDGEFLRTSCGSPRYAPAEYISGLYAGP 192  
Qy 210 RADMMSCGVLFFALLVGLALPPDDNLRQLLEKVRGVFHMFFIPDCOSLLRGMEYEP 269  
Db 193 EVDIMSSGVLIALLCGLTFPDDHVPFLFKKIDGIFTYQYINPVSISLKHMLQYDP 252  
Qy 270 EKRLSLEQIOKHPWYLGKHEPDPCLPAPGRVAMRSLSNGELDPVLESMAISLGCGR 330  
Db 253 MKRATIKDIREHEWF---KQDLPSYLFPE-----DPSYSSTMI----- 286  
Qy 331 DRELRHRELSEENOEKMI-----YYLLDRKERYPCGEODL---PP 371  
Db 287 DDEALKEVCEKFCSESEVLSCLYNRHQDPLAVAYHLIDNRIRIMNAKDFYATSP 345

Db 303 TESEVNNISYSGPDQDLAVAYHLIDNRIRIMNAKDFYATSP 362  
Qy 387 LSRHGRKP 395  
Db 363 LKPHRMP 371

RESULT 12  
US-09-101-146-6  
; Sequence 6, Application US/09101146  
; Patent No. 6124125  
; GENERAL INFORMATION:  
; APPLICANT: Dartmouth College, St. Vincents Institute of  
; APPLICANT: Medical Research, Kemp et al.  
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jane Massey Licata, Esq.  
; STREET: 66 E. Main Street  
; CITY: Marlton  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE  
; COMPUTER: IBM PC  
; OPERATING SYSTEM: WINDOWS 95  
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/101,146  
; FILING DATE: October 7, 1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PN7450  
; FILING DATE: 8 JAN 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jane Massey Licata  
; REGISTRATION NUMBER: 32,257  
; REFERENCE/DOCKET NUMBER: DC-0050  
; TELEPHONE: (856) 810-1515  
; TELEFAX: (856) 810-1454  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 257  
; TYPE: Amino acid  
; TOPOLOGY: Linear  
US-09-101-146-6

Query Match 17.1%; Score 700.5; DB 3; Length 257;  
Best Local Similarity 50.6%; Pred. No. 1.5e-37;  
Matches 129; Conservative 52; Mismatches 73; Indels 1; Gaps 1;

Qy 32 GYRLKLTGKGTGGLVLCVHCITGQKVAIKIYVREKL-SESYLKVREITAIKLE 90  
Db 1 GYRLKLTGKGTGGLVLCVHCITGQKVAIKIYVREKL-SESYLKVREITAIKLE 90  
Qy 91 HPHVYLVGDTLVGTFGVKKGHEHDTGKVAIKIYVREKL-SESYLKVREITAIKLE 149  
Db 61 HPHVYLVGDTLVGTFGVKKGHEHDTGKVAIKIYVREKL-SESYLKVREITAIKLE 132  
Qy 150 YSICHLRLKPNLILDEKNNIRIDAFGNASLOVGSLSLETSCGSPHYACPEYISGLYAGP 209  
Db 133 HMYVHRLDKPENVLLDAMNNAKIDAFGLSNMMSDGEFLRTSCGSPRYAPAEYISGLYAGP 192  
Qy 210 RADMMSCGVLFFALLVGLALPPDDNLRQLLEKVRGVFHMFFIPDCOSLLRGMEYEP 269  
Db 193 EVDIMSSGVLIALLCGLTFPDDHVPFLFKKIDGIFTYQYINPVSISLKHMLQYDP 252  
Qy 270 EKRLSLEQIOKHPWYLGKHEPDPCLPAPGRVAMRSLSNGELDPVLESMAISLGCGR 330  
Db 253 MKRATIKDIREHEWF---KQDLPSYLFPE-----DPSYSSTMI----- 286  
Qy 331 DRELRHRELSEENOEKMI-----YYLLDRKERYPCGEODL---PP 371  
Db 287 DDEALKEVCEKFCSESEVLSCLYNRHQDPLAVAYHLIDNRIRIMNAKDFYATSP 345



```

Db      120 KIVHRDIAKPEHLLDDEHINVAKIADPGLSNIMFDGNEFLKTSKSGSPNYAAPEVLSKLIYAGP 179
QY      211 RADMMSCGVILLFALLVGLALPEPDDNLRQLLEKVKRGVFMHPHFIPPDCCSLRGMIIEVP 270
Db      180 EVDVWSCGVILLYVLCRILPEPDDSEIPVLFKNINSNGVYTLPLKFLSPGAAGLIKRMILYVP 239
QY      271 EKRLSEIQIOKHFWY 285
Db      240 LNRISIHIMDDWF 254

```

## RESULT 15

```

US-09-930-181-18
; Sequence 18, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 160 101 V1
; CURRENT APPLICATION NUMBER: US/09/930.181
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 18
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-181-18

```

```

Query Match      14.1%; Score 579; DB 4; Length 149;
Best Local Similarity 80.7%; Pred. No. 4.2e-30;
Matches 117; Conservative 13; Mismatches 7; Indels 8; Gaps 2;

```

```

QY      579 MSSTPSSPELAKRSWGFNFSIDKEQIFLYLKDPLSSIKADIVHAFSLIPSLSHSV 638
Db      1 MSNITPSSPELAKRSWGFNFSIDKEQIFVYIKDKPLSSIKADIVHAFSLIPSLSHSV 60
QY      639 LSQTSFRAEYKASGSPSVFQKPVRRQVDISSSEGEPEPSPRRDGGGCIYSVFTLLISGP 698
Db      61 ISQTSFRAEYKATGPAVFQKPKFQVDITYTEGGEAQKE-----NGIYSVFTLLISGP 114
QY      699 SRRKRVETIQAOQLSTHD--QPS 721
Db      115 SRRKRVETIQAOQLSTHDPLRPS 139

```

Search completed: April 16, 2003, 12:33:27  
Job time : 24 secs

100

100

100

GenCore version 5.1.4\_P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 12:28:03 ; Search time 98 Seconds  
(without alignments)  
1635.762 Million cell updates/sec

Title: US-10-003-690-2  
Perfect score: 4108  
Sequence: 1 MSGAKGEGGSPAYHLPHP.....PRRGPPKKLLATNGTLP 778

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL.21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4108	100.0	778	4	08TDC2	08TDC2 homo sapien
2	3956.5	96.3	794	4	08TDC3	08TDC3 homo sapien
3	3758	91.5	715	4	096J14	096J14 homo sapien
4	2490.5	60.6	603	4	060843	060843 homo sapien
5	2466	47.5	851	5	09AV4	09AV4 homo sapien
6	1952	42.5	914	5	061298	061298 caenorhabdi
7	1814	44.2	698	5	019469	019469 drosophila
8	1744	42.5	914	5	095T82	095T82 drosophila
9	1228	29.9	701	5	08SSX5	08SSX5 dictyosteli
10	874	21.3	833	5	09Y2K2	09Y2K2 homo sapien
11	832	20.3	1371	4	096RG0	096RG0 homo sapien
12	828	20.2	691	4	015524	015524 homo sapien
13	827	20.1	755	4	096HB3	096HB3 homo sapien
14	827	20.1	755	4	096YD8	096YD8 homo sapien
15	826	20.1	689	4	096JG7	096JG7 homo sapien
16	823	20.0	689	4	096JG7	096JG7 homo sapien

17	823	20.0	752	4	096L34	096L34 homo sapien
18	822	20.0	722	11	008679	008679 rattus norv
19	814	19.8	1192	5	09TW45	09TW45 caenorhabdi
20	814	19.8	1192	5	017346	017346 xenopus lae
21	812.5	19.7	1096	5	017368	017368 caenorhabdi
22	809	19.6	795	4	09P0L2	09P0L2 homo sapien
23	806.5	19.6	729	4	08TBA1	08TBA1 homo sapien
24	796	19.4	535	10	08RWD2	08RWD2 arabidopsi
25	795.5	19.4	729	4	060219	060219 homo sapien
26	795	19.3	401	8	098SC8	098SC8 guillardi
27	793	19.3	504	10	093113	093113 cucumis sat
28	792.5	19.3	744	4	096RG1	096RG1 homo sapien
29	792	19.3	729	11	09JKE4	09JKE4 mus musculu
30	789	19.2	729	11	09JKE4	09JKE4 mus musculu
31	789	19.2	926	4	09HOK1	09HOK1 mus musculu
32	788.5	19.2	744	11	09JKE5	09JKE5 mus musculu
33	788	19.2	753	4	08W83	08W83 mus musculu
34	788	19.2	793	11	008678	008678 rattus norv
35	786	19.1	514	10	09X25	09X25 mus musculu
36	784	19.1	795	11	08VH5	08VH5 mus musculu
37	783.5	19.1	514	10	09W726	09W726 lycopersico
38	781	19.0	797	11	08VH5	08VH5 mus musculu
39	780.5	19.0	798	13	091A88	091A88 gallus gall
40	777	18.9	505	10	09ZRL1	09ZRL1 oryza sativ
41	776	18.9	938	5	09Y8V8	09Y8V8 drosophila
42	772	18.8	1075	5	095U75	095U75 drosophila
43	770.5	18.8	752	4	09UN34	09UN34 homo sapien
44	767	18.7	1398	5	077268	077268 drosophila
45	767	18.7	1398	5	09W532	09W532 drosophila

## ALIGNMENTS

RESULT 1	PRELIMINARY:	PRT:	778 AA.
08TDC2			
1	08TDC2		
AC	01-JUN-2002 (TREMBLrel. 21, Created)		
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Protein kinase-like protein.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID-9606;		
RN	SEQUENCE FROM N.A.		
RP	She X.Y., Yu L., Guo J.H.;		
RA	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.		
RL	EMBL: AF479827; AAL87698.1;		
DR	EMBL: AF479827; AAL87698.1;		
KM	SEQUENCE 778 AA; 85086 MW; 8D1818DA54398BB CRC64;		
SO			
Query Match	100.0%;	Score 4108;	DB 4; Length 778;
Best Local Similarity	100.0%;	Pred. No. 7e-264;	
Matches 778; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;

OY	1	MSGAKGEGGSPAYHLPHPHPPOHAYGVPYRLKTLGKGTGLVLYGHCITGQKV	60
DB	1	MSGAKGEGGSPAYHLPHPHPPOHAYGVPYRLKTLGKGTGLVLYGHCITGQKV	60
OY	61	AIKIVREKLSVLMKVERETAIKLILHPVYLKLDHYENKKYLYLLEHVSGETLPD	120
DB	61	AIKIVREKLSVLMKVERETAIKLILHPVYLKLDHYENKKYLYLLEHVSGETLPD	120
OY	121	VLVKKRGLTPKPKRKFROIVSALDFCHSYICHRDLKPENLDDKNNRIADFQMASL	180
DB	121	VLVKKRGLTPKPKRKFROIVSALDFCHSYICHRDLKPENLDDKNNRIADFQMASL	180
OY	121	VLVKKRGLTPKPKRKFROIVSALDFCHSYICHRDLKPENLDDKNNRIADFQMASL	180
DB	121	VLVKKRGLTPKPKRKFROIVSALDFCHSYICHRDLKPENLDDKNNRIADFQMASL	180
OY	181	QVDSILLETSCGSPHYACPEVIRGEKYDGRADMSGCVILFALVGCALPFDDDNRL	240
DB	181	QVDSILLETSCGSPHYACPEVIRGEKYDGRADMSGCVILFALVGCALPFDDDNRL	240

```

QY 241 EKVYKGVFHMHPFIIPDCQSLIRGMIEVEPEKRLSLBOIOKHPWYLGKHEPDCLEPAP 300
DB 241 EKVYKGVFHMHPFIIPDCQSLIRGMIEVEPEKRLSLBOIOKHPWYLGKHEPDCLEPAP 300
QY 301 GRRVAMSLNSNGELDDVLESMAISLCCFDRERLHRELSEENOEKMIYLLDRKER 360
DB 301 GRRVAMSLNSNGELDDVLESMAISLCCFDRERLHRELSEENOEKMIYLLDRKER 360
QY 361 YPSCEDODLPPRNDVDPKRRVDSVPMLSRHKRPERKSMENVLSTIDAGGGSPVPTTRA 420
DB 361 YPSCEDODLPPRNDVDPKRRVDSVPMLSRHKRPERKSMENVLSTIDAGGGSPVPTTRA 420
QY 421 LEMAHOSRBSYSGASTGSSPSPSPVFSFSPGAGDEARGGSPSTKTOTLPS 480
DB 421 LEMAHOSRBSYSGASTGSSPSPSPVFSFSPGAGDEARGGSPSTKTOTLPS 480
QY 481 RSPRGAGAGEOPPPPARSTPLPGPPSPSSGCTPLHSPHTPRASPTGTPTTPTTPTT 540
DB 481 RSPRGAGAGEOPPPPARSTPLPGPPSPSSGCTPLHSPHTPRASPTGTPTTPTTPTT 540
QY 541 GGGVGAAMRSKLSIRNSFLGSPRHRKMOVPTAEEMSLTPESSPELAKRSMGNT 600
DB 541 GGGVGAAMRSKLSIRNSFLGSPRHRKMOVPTAEEMSLTPESSPELAKRSMGNT 600
QY 601 SLDKREQIFLVKDKPLSLIKADIVHAFLSTPSLSHVSLSQTSFRAEYASGSPVFOQP 660
DB 601 SLDKREQIFLVKDKPLSLIKADIVHAFLSTPSLSHVSLSQTSFRAEYASGSPVFOQP 660
QY 661 VRFVDISSLSEPPSPRDSGGGIGYVTFLLISGSRKRVETIOAQLLSTHDP 720
DB 661 VRFVDISSLSEPPSPRDSGGGIGYVTFLLISGSRKRVETIOAQLLSTHDP 720
QY 721 SVQALADEKNGAOTRPAAPRSLQPPGRDPELSSPPRGGPKKLATNGTPLP 778
DB 721 SVQALADEKNGAOTRPAAPRSLQPPGRDPELSSPPRGGPKKLATNGTPLP 778

```

## RESULT 2

```

ID 08TDC3 PRELIMINARY; PRT; 794 AA.
AC 08TDC3;
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Putative serine/threonine protein kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA She X.Y., Guo J.H., Yu L.;
RT Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF479826; AAL87697.1;
KW Kinase; Serine/threonine-protein kinase.
SQ SEQUENCE 794 AA; 86753 MW; 5DD395B0E61AEF77 CRC64;

```

Query Match 96.3%; Score 3956.5; DB 4; Length 794;

Best Local Similarity 98.1%; Pred. No. 7.8e-254;

Matches 756; Conservative 1; Mismatches 9; Indels 5; Gaps 1;

```

QY 8 GGGSPAYHLPRHPHPOHAYGPRYRLEKTLGKQGLVNLGVHCITGQVAKIYVR 67
DB 29 GGGG-----EEAEERGHAYGPRYRLEKTLGKQGLVNLGVHCITGQVAKIYVR 83
QY 68 EKLSSVLMKVERELAILKLEHPVYLKLDHYENKKYLLYLEHVSGLPFDYVKKGR 127
DB 84 EKLSSVLMKVERELAILKLEHPVYLKLDHYENKKYLLYLEHVSGLPFDYVKKGR 143
QY 128 LTPKARKEFFQIYSALDFCHSYISCHDLPENLLDEKNINRIADGMSLQVGSLL 187
DB 128 LTPKARKEFFQIYSALDFCHSYISCHDLPENLLDEKNINRIADGMSLQVGSLL 187

```

```

DB 144 LTPKARKEFFQIYSALDFCHSYISCHDLPENLLDEKNINRIADGMSLQVGSLL 203
QY 188 ETCGSPYACPEYIKGKYGRADWMSGVILFALLVGLAPEDDMLROLLEKVRGV 247
DB 204 ETCGSPYACPEYIKGKYGRADWMSGVILFALLVGLAPEDDMLROLLEKVRGV 263
QY 248 FHMHPFIIPDCQSLIRGMIEVEPEKRLSLBOIOKHPWYLGKHEPDCLEPAPRVA 307
DB 264 FHMHPFIIPDCQSLIRGMIEVEPEKRLSLBOIOKHPWYLGKHEPDCLEPAPRVA 323
QY 308 SLPSNGELDDVLESMAISLCCFDRERLHRELSEENOEKMIYLLDRKERPSCEDQ 367
DB 324 SLPSNGELDDVLESMAISLCCFDRERLHRELSEENOEKMIYLLDRKERPSCEDQ 383
QY 368 DLPPRNDVDPKRRVDSVPMLSRHKRPERKSMENVLSTIDAGGGSPVPTTRA 427
DB 384 DLPPRNDVDPKRRVDSVPMLSRHKRPERKSMENVLSTIDAGGGSPVPTTRA 443
QY 428 QRSRVSAGASTGLSSPSPSPVFSFSPGAGDEARGGSPSTKTOTLPSRGRGG 487
DB 444 QRSRVSAGASTGLSSPSPSPVFSFSPGAGDEARGGSPSTKTOTLPSRGRGG 503
QY 488 AGROPPEPARSTPLPGPPSPSSGCTPLHSPHTPRASPTGTPTTPTTPTT 547
DB 504 AGROPPEPARSTPLPGPPSPSSGCTPLHSPHTPRASPTGTPTTPTTPTT 563
QY 548 AMRSKLSIRNSFLGSPRHRKMOVPTAEEMSLTPESSPELAKRSMGNT 607
DB 564 AMRSKLSIRNSFLGSPRHRKMOVPTAEEMSLTPESSPELAKRSMGNT 623
QY 608 IFLVLDKRLSLIKADIVHAFLSTPSLSHVSLSQTSFRAEYASGSPVFOQP 667
DB 624 IFLVLDKRLSLIKADIVHAFLSTPSLSHVSLSQTSFRAEYASGSPVFOQP 683
QY 668 SSSSEPPSPRDSGGGIGYVTFLLISGSRKRVETIOAQLLSTHDP 727
DB 684 SSSSEPPSPRDSGGGIGYVTFLLISGSRKRVETIOAQLLSTHDP 743
QY 728 EKNGAOTRPAAPRSLQPPGRDPELSSPPRGGPKKLATNGTPLP 778
DB 744 EKNGAOTRPAAPRSLQPPGRDPELSSPPRGGPKKLATNGTPLP 794

```

## RESULT 3

```

ID 096JL4 PRELIMINARY; PRT; 715 AA.
AC 096JL4;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE KIAA1811 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA MEDLINE-21425130; PubMed-11347906;
RA Nagase T., Nakayama M., Nakajima D., Kitano R., Ohara O.;
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.
RL DNA Res. 8:85-95(2001).
DR EMBL; AB058714; BAB47440.1;
DR InterPro; IPR000719; Euk_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR pfam; PF00069; kinase_1.
DR Prodom; PD00001; Euk_kinase_1.
DR PROSITE; PS5011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS0108; PROTEIN_KINASE_ST; UNKNOWN_1.

```

FT NON\_TER 1 1  
 SQ SEQUENCE 715 AA; 78499 MW; B906EE115C418A5 CRC64;  
 Query Match  
 Best Local Similarity 100.0%; Score 3758; DB 4; Length 715;  
 Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

64 IYNRKLSSEYLVKVERIALTKLIEHPHYKLHDYENKKYLYLVEHVSGLDFDYL 123  
 1 IYNRKLSSEYLVKVERIALTKLIEHPHYKLHDYENKKYLYLVEHVSGLDFDYL 60  
 124 KKGRLTPKARKKFFQIYVADFCCHSYISICHRDLKPEMLLDKNNIRIADFGMAISLVG 183  
 61 KKGRLTPKARKKFFQIYVADFCCHSYISICHRDLKPEMLLDKNNIRIADFGMAISLVG 120  
 184 DSLLETSGSPHYACPEVIRKGEKYDGRADMMSCVILFALLVGLPPDDNLRLQLERK 243  
 121 DSLLETSGSPHYACPEVIRKGEKYDGRADMMSCVILFALLVGLPPDDNLRLQLERK 180  
 244 KKGRLTPKARKKFFQIYVADFCCHSYISICHRDLKPEMLLDKNNIRIADFGMAISLVG 303  
 181 KKGRLTPKARKKFFQIYVADFCCHSYISICHRDLKPEMLLDKNNIRIADFGMAISLVG 240  
 304 VAMRSLPSNGELDPVLESMASLGCPRERLHRELSEENOEKMIYLLDRKERYPS 363  
 241 VAMRSLPSNGELDPVLESMASLGCPRERLHRELSEENOEKMIYLLDRKERYPS 300  
 364 CEDODLPKRDVDPKRVDSPLSHRGRKRPERSMEVLSITDAGGSPVPTRALEM 423  
 301 CEDODLPKRDVDPKRVDSPLSHRGRKRPERSMEVLSITDAGGSPVPTRALEM 360  
 424 AHSQRSRYSVSGASTGLSSPLSPSPVSPSPSPSPSPSPSPSPSPSPSPSPSPSP 483  
 361 AHSQRSRYSVSGASTGLSSPLSPSPSPVSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420  
 484 RGGAGAGOPPPPSARSTPLPGPSPSPSPSGCTPLHSLATPRASPTGTPPTPPSPGG 543  
 421 RGGAGAGOPPPPSARSTPLPGPSPSPSPSGCTPLHSLATPRASPTGTPPTPPSPGG 480  
 544 VGGAMRSRLNLSINSLGSPFRHRKMOVTAEMSLPESPSPSPSPSPSPSPSPSPSPSP 603  
 481 VGGAMRSRLNLSINSLGSPFRHRKMOVTAEMSLPESPSPSPSPSPSPSPSPSPSPSP 540  
 604 KEEQIFLVKPKPLSSIKADIYHAFSLPSLSHVSLSQTSFRAEYKASGGSVQKPPRF 663  
 541 KEEQIFLVKPKPLSSIKADIYHAFSLPSLSHVSLSQTSFRAEYKASGGSVQKPPRF 600  
 664 QVDISSEGPSPRRDGGGGGIGTFTLISGSRFRKRVETIOAQLSTHDQPSVQ 723  
 601 QVDISSEGPSPRRDGGGGGIGTFTLISGSRFRKRVETIOAQLSTHDQPSVQ 660  
 724 ALADKNGAQTPAGAPPSLQPPGRDPPELSSSPRRQPPKDKLLATNGTLP 778  
 661 ALADKNGAQTPAGAPPSLQPPGRDPPELSSSPRRQPPKDKLLATNGTLP 715

RESULT 4  
 060843 PRELIMINARY: PRT: 603 AA.

AC 060843;  
 DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)  
 DT 01-AUG-1998 (T-EMBLrel. 20, Last annotation update)  
 DE Putative serine/threonine protein kinase (Fragment).  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCBI\_TaxID=9606;  
 RX MEDLINE=21064499; PubMed=11124703;  
 RA Stanchi F., Bertocco E., Topo S., Dioguardi R., Simonati B.,

RA Cannata N., Zimbello R., Lanfranchi G., Valle G.; high similarity to  
 "Characterization of 16 novel human genes showing high similarity to  
 yeast sequences".  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: AJ006701; CA007196.1; --  
 DR HSSP: O63450; 1A06.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR InterPro: IPR01245; Tyr\_kinase.  
 DR Pfam: PF00069; pkinase.1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR PRODOM: PD000001; Euk\_pkinase.1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM.1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_SF.1.  
 KW Kinase; Serine/threonine-protein kinase.

FT NON\_TER 1  
 SQ SEQUENCE 603 AA; 67401 MW; B02C5D678F8E96 CRC64;

Query Match  
 Best Local Similarity 72.8%; Score 2490.5; DB 4; Length 603;  
 Matches 488; Conservative 52; Mismatches 45; Indels 85; Gaps 9;

87 LIEHPHYKLHDYENKKYLYLVEHVSGLDFDYLKKGRLTPKARKKFFQIYVADFC 146  
 1 LIEHPHYKLHDYENKKYLYLVEHVSGLDFDYLKKGRLTPKARKKFFQIYVADFC 60  
 147 CHSYISICHRDLKPEMLLDKNNIRIADFGMAISLVGSLLETSGSPHYACPEVIRK 206  
 61 CHSYISICHRDLKPEMLLDKNNIRIADFGMAISLVGSLLETSGSPHYACPEVIRK 120  
 207 YGRRADMMSCVILFALLVGLPPDDNLRLQLERKRVGFHMFHTIPDCQSLIRMI 266  
 121 YGRRADMMSCVILFALLVGLPPDDNLRLQLERKRVGFHMFHTIPDCQSLIRMI 180  
 267 EYEPERKLSLEQIYVADFCCHSYISICHRDLKPEMLLDKNNIRIADFGMAISLVG 326  
 181 EYEPERKLSLEQIYVADFCCHSYISICHRDLKPEMLLDKNNIRIADFGMAISLVG 238  
 327 GCFRDERLHRELSEENOEKMIYLLDRKERYPSCEDOLPPRDVDPKRVDSM 386  
 239 GCFRDERLHRELSEENOEKMIYLLDRKERYPSCEDOLPPRDVDPKRVDSM 298  
 387 LSRHGRKRPERSMEVLSITDAGGSPVPTRALEMAHSQRSRYSVSGASTGLSSPLS 446  
 299 LSRHGRKRPERSMEVLSITDAGGSPVPTRALEMAHSQRSRYSVSGASTGLSSPLS 355  
 447 SPRSPVSPSPSPSPSGCTPLHSLATPRASPTGTPPTPPSPGG 506  
 356 SPRSPVSPSPSPSPSGCTPLHSLATPRASPTGTPPTPPSPGG 374  
 507 GSPRSSGCTPLHSLATPRASPTGTPPTPPSPGGVGGAMRSRLNLSINSLGSPFR 566  
 375 GSPRSSGCTPLHSLATPRASPTGTPPTPPSPGGVGGAMRSRLNLSINSLGSPFR 421  
 567 HRRKMOVTAEMSLPESPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 626  
 422 HRRKMOVTAEMSLPESPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 481  
 627 AFLSIPSLSHVLSQTSFRAEYKASGGSVQKPPRQVODISSSGPSPRRDGGGG 686  
 482 AFLSIPSLSHVLSQTSFRAEYKASGGSVQKPPRQVODISSSGPSPRRDGGGG 535  
 687 IYVTFITLISGSRFRKRVETIOAQLSTHDQPSVQALADKNGAQTPAGAPPSLQ 746  
 536 IYVTFITLISGSRFRKRVETIOAQLSTHDQPSVQALADKNGAQTPAGAPPSLQ 481  
 747 PPGRPPELS 756  
 578 PPGRPPELS 585

RESULT 5

096AV4 ID 096AV4 PRELIMINARY; PRT; 473 AA.  
AC 096AV4;  
DT 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Hypothetical 50.8 kDa protein.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN NCBI\_TaxID=9606;  
RP SEQUENCE FROM N.A.  
RC TISSUE=LYMPH;  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC016681; AAL16681.1;  
DR InterPro: IPR000449; UBA\_domain.  
KW Hypothetical protein.  
SQ SEQUENCE 473 AA: 50750 MW: D27DAD437DA96AA9 CRC64;

Query Match  
Best Local Similarity 60.0%; Score 2466; DB 4; Length 473;  
Matches 472; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

0Y 306 MRSLSNGELDPVLESMASLCGCFDRERLRHRESEENDEKMTYLLDRKERYPSC 365  
1 MRSLSNGELDPVLESMASLCGCFDRERLRHRESEENDEKMTYLLDRKERYPSC 60  
DB 1 MRSLSNGELDPVLESMASLCGCFDRERLRHRESEENDEKMTYLLDRKERYPSC 60  
0Y 366 DODLPKRDVPPPKRRVDSPLMSLRHRRPKRMEVLSITDAGGSGPPTRALEMAQ 425  
61 DODLPKRDVPPPKRRVDSPLMSLRHRRPKRMEVLSITDAGGSGPPTRALEMAQ 120  
DB 426 HSORSRSVSGASTGLSSPLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 485  
121 HSORSRSVSGASTGLSSPLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 180  
0Y 486 GGAEPDPPPPPARSTPLPGPPGSRSSGTPPLHSPLTPASPTGPTPPSPGGG 545  
181 GGAEPDPPPPPARSTPLPGPPGSRSSGTPPLHSPLTPASPTGPTPPSPGGG 240  
DB 546 GAAMRSRLNSIRNSFLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 605  
241 GAAMRSRLNSIRNSFLSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 300  
0Y 606 EOLFLVKKPKLSSIKADIVHAFSLPSLSHVSLSQTSFRAEYKASGGSVOKPVRQY 665  
301 EOLFLVKKPKLSSIKADIVHAFSLPSLSHVSLSQTSFRAEYKASGGSVOKPVRQY 360  
DB 666 DISSSGPSPRRDGGGGGTYVTFTLISGSPRRFRKAVETIOAQLSTHDQSVQAL 725  
361 DISSSGPSPRRDGGGGGTYVTFTLISGSPRRFRKAVETIOAQLSTHDQSVQAL 420  
0Y 726 ADEKGAQTRPAGAPRSLQPPGGRDDELSSPPRRGPKDKKLATNGTPTP 778  
421 ADEKGAQTRPAGAPRSLQPPGGRDDELSSPPRRGPKDKKLATNGTPTP 473

RESULT 6  
09VUVA  
09VUVA PRELIMINARY; PRT; 851 AA.  
AC 09VUVA;  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
DE CG6114 protein.  
GN CG6114.  
OS Drosophila melanogaster (fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Phlebotomidae; Diptera; Drosophila.  
RN NCBI\_TaxID=7227;  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.C., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu L., Cantler A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke A., Denger Z., Mays A.D., Dew I., Dietz S.M.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heilmann T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jallat M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lai Y., Levitsky A.A., Li Z., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Misha N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,  
RA Palazzolo M., Pittman K.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden Klamas I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spindling A.C., Stapleton M., Strong R., Sun E.,  
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Weissman D.A., Weissman G.M., Weissman J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
CC -1. SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AE003529; AAF9569.1;  
DR HSSP; Q63450; 1A06.  
DR FLYbase; FB90036544; CG6114.  
DR InterPro: IPR000719; Euk\_Pkinase.  
DR InterPro: IPR002290; Ser-Thr\_Pkinase.  
DR InterPro: IPR001245; Tyr\_Pkinase.  
DR Pfam; PF00069; Pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR PRODOM; PD000001; Euk\_Pkinase; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 851 AA: 94152 MW: 9E5456B8F5D47E0 CRC64;

Query Match  
Best Local Similarity 54.9%; Score 1952; DB 5; Length 851;  
Matches 421; Conservative 96; Mismatches 138; Indels 112; Gaps 24;

0Y 26 QHQQVGPRLKLTCKGOTGLVKLGCHITGQKVAIKIVNEKLSSEYLMKVEREIAL 85  
10 ENQGVGPRLKLTCKGOTGLVKLGCHITGQKVAIKIVNEKLSSEYLMKVEREIAL 69  
DB 86 KLEHPVLTLDVYENKRYLYLVLEHVSGLLEFDYLVKGLPKRKARKEFRQVSLAD 145  
70 KLEHPVLTLDVYENKRYLYLVLEHVSGLLEFDYLVKGLPKRKARKEFRQVSLAD 129  
0Y 146 FCHSYSTCHRLKPNLLIDKNNRIADFGMASLOVDSLETSQSGSPHACPEVIRGE 205  
130 FCHSYSTCHRLKPNLLIDKNNRIADFGMASLOVDSLETSQSGSPHACPEVIRGE 189

QY 206 KYDGRADWMSGVILFALLVGLALPDDNLRKLEKRGVFNHPRFPDCCSLRGM 265  
 DB 190 KYDGRADWMSGVILFALLVGLALPDDNLRKLEKRGVFNHPRFPDCCSLRGM 249  
 QY 266 IEVEPERKLSLEIOKHPWY-LOGKHEPD---PCLPAPGRVAMSLPSNGELDPVLE 321  
 DB 250 IEVAPRRRLTAEINRHPWYAGGKGBLELPMMEV-----VQTHVPTAVAPVDVLAN 304  
 QY 322 SMAALGCEPRERLRLRELSSENOEKMIYLLDKERKPSCEDD---LPPRD---V 375  
 DB 305 AICISGCEKEKELIOLSSNTEKVIYELLERRRRPALEDDDELAKSRSLDAV 364  
 QY 376 DPPRRKVDSPMLSRHGRKRRPKRSMELSTIDAGGGSPY-PTRRALMAQHS---QHS 430  
 DB 365 DPPRRRLDTCRI--NGTAPSYGOL-----SESGPLTPRQAFNFRSYSTNRHOR 413  
 QY 431 RVSAGSLGSLSSPLSP--RSVPFS-----ESPEPA-----GDEARGG-GS 470  
 DB 414 RSPTYVTSVRSSTSHPTRCNSPMSAQQOAMASRPSPACTRSTYGDNRSGHNS 473  
 QY 471 PYSKTOTLPS-----RGRGGGAGEQPPPSASTPLPG-PPG 507  
 DB 474 SVGRFSSHSSOKSIGEDVYVVRPREIRERDLSKORGG-----SPRDGDCGLPPG 525  
 QY 508 SP--RSAGGTPHSPH-----TPRASP---TGPGTTPPSGGGGAAM 549  
 DB 526 SPGGSSGTSASPSVHHRANSPTAIIVNPGSPMANNSSPGM--PGSCNTPGGGLW 583  
 QY 550 RSLRNSRNSFLGSPRRHRKMOVPFAEMSSLTPSSPELAKRWFNGFNISLDEKQIF 609  
 DB 584 KTRLTNKNLSLSPRRHRKMOV-SADEV-FLTPSSPELKRSMFGMLITTEDEFT 641  
 QY 610 IYLDKPLSLIKADIVAFSLSPSLSHSVLSQTSFRAFKASG-GPSVQKQVRECVDS 668  
 DB 642 ILVKKPLATVAKHLIHLFSLMAELSHSVSPISFVEKFKRNGCVPMORHKKFQVDS 701  
 QY 669 SSGEPSPRRDGGSGGIVSVFTLLSGSPRRKRVETIOALIS 715  
 DB 702 AI-----CKGDIADMLFALTLLSGNIRRRICEHIQSVCS 741  
 RESULT 7  
 ID 061298 PRELIMINARY; PRT: 698 AA.  
 AC 061298: TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
 DE HPOK-1 protein.  
 GN HPOK-1.  
 OS Halocynthia roretzi (Sea squirt).  
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
 OC Stolidobranchia; Pyuridae; Halocynthia.  
 OC NCBI\_TaxID=7729;  
 RN  
 RP  
 RC  
 RX MEDLINE=98440280; PubMed=9767157;  
 SA Sasaki Y., Ogasawara M., Makabe K.W.:  
 RT "Maternally localized RNA encoding a serine/threonine protein kinase  
 in the ascidian, Halocynthia roretzi."  
 RL Mech. Dev. 76:161-163(1998).  
 CC 1. SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: AB014885; BAA28663.1.  
 DR HSSP: Q63450.1A06  
 DR InterPro: IPR000719; Euk\_Pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_Pkinase.  
 DR Pfam: PF00069; Pkinase; 1.  
 DR ProDom: PD000001; Euk\_Pkinase; 1.  
 DR SMART: SM00220; S\_TKC; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.

DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Serine/threonine-protein kinase; transferase.  
 SO SEQUENCE 698 AA; 78607 MW; B29B1751E8300BF CRC64;  
 Query Match 44.2%; Score 1814; DB 5; Length 698;  
 Best local similarity 52.8%; Pred. No. 5,6e-112;  
 Matches 383; Conservative 109; Mismatches 115; Indels 118; Gaps 21;  
 QY 25 POHAQVYPRLEKTIKGTGOTGVKGVHCITGOKAIKIVREKLSSVLMKVERELAI 84  
 DB 5 POPQVYPRLEKTIKGTGOTGVKGVHCITGOKAIKIVREKLSSVLMKVERELAI 64  
 QY 85 LKLEHPRVLAHVDYENKRYLYLVLEHVSGLPFLYTKKRRPKARKEFPQIVSL 144  
 DB 65 KMLEHPRVLAHVDYENKRYLYLVLEHVSGLPFLYTKKRRPKARKEFPQIVSL 124  
 QY 145 DECHSYSTCHDLKPEMLILDEKNNIRIADFGMASLQVGDLSLETSCGPHVACPEYVIG 204  
 DB 125 DYCHNNHVCCHDLKPEMLILDEKNNIRIADFGMASLQVGDLSLETSCGPHVACPEYVIG 184  
 QY 205 KYDGRADWMSGVILFALLVGLALPDDNLRKLEKRGVFNHPRFPDCCSLRGM 264  
 DB 185 EYDGRADWMSGVILFALLVGLALPDDNLRKLEKRGVFNHPRFPDCCSLRGM 244  
 QY 265 MIEVEPERKLSLEIOKHPWYLOGKHE-----PD---PCLPAPGRVAMSLPSNGE 314  
 DB 245 MIDVPRDKRLSLQOVLQHPWYAGGKGBLELPMMEV-----VQTHVPTAVAPVDVLAN 304  
 QY 315 LDDVLESMAISLGCFFDRERLRLRELSSENOEKMIYLLDKERKPSCEDD---LPPRN 373  
 DB 297 VDDVLAWSLGCFFDRERLRLRELSSENOEKMIYLLDKERKPSCEDD---LPPRN 356  
 QY 374 DVDPKRRVD--SPMLSRHGRKRRPKRSMELSTIDAGGGSPYPTRRALMAQHS 425  
 DB 357 HPDPRKRRVSTSLSSNDGCVNPIQKMSASLITLSS--SSPLSRK--KSE 411  
 QY 426 HSQRSVYSGASTGLSSPLSPSPSVFSPSPGAGDEARGGSPSTKYOTLPSRPG 485  
 DB 412 THORSOSL--TESSNRL-----VCNISDQTKESKRNGTTPVRKQTCSS----- 456  
 QY 486 GGAGGPPPPSARSTPLPGPPSPSSGGTPLHSPHTPASPSTGTPPTPPSPGGVAG 545  
 DB 457 ---NQPVP-----QINTPAS-----FNP----- 471  
 QY 546 GAAMSRLNLSNLSGSPRRHRKMOVPFAEMSSLTPSSPELAKRWFNGFNISLDEKQIF 609  
 DB 472 ---WRQRLASLNTKMGSPRRHRKMOVPFAEMSSLTPSSPELAKRWFNGFNISLDEKQIF 609  
 QY 604 ---KEQIFLYLDKPLSLIKADIVAFSLSPSLSHSVLSQTSFRAFKASG-GPSVQKQVRECVDS 668  
 DB 528 STEHCEDELPAIVAKKRLNLSVKSSELVAFSLPILTHSMVSPISFVEKFKRNGCVPMORHKKFQVDS 701  
 QY 657 F-QKPVREFOVDI-----SSSGEPSPRRDGGSGGIVSVFTLLSGSPRRKRVETIOALIS 715  
 DB 588 FHGRSIFQVDDIIQHSLSLQDENGKKRPSQYGS-----FTIAFSLISGPIRRYRVLEL 642  
 QY 709 IOAQL 713  
 DB 643 LOMOM 647  
 RESULT 8  
 ID 019469 PRELIMINARY; PRT: 914 AA.  
 AC 019469: OGBM6;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
 DE F15A2.6 protein (Serine/threonine kinase SMD-1).  
 GN F15A2.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Rhabditinae; Caenorhabditidae.

QY 481 -----RGRGGAGE-----QPPPSANSTPLPGPPSPRS 511  
 Db 495 AARHVDAOERERESNDGRGSRKESKDRSKASSSCKNDASTSSVPHKSPSPVMS 554  
 QY 512 SGGTPLHSLPLTPRASPGTGTTPPPPGGCGVGAAMRSLNLSIRNLSLSPFFHRRKM 571  
 Db 555 ESYVVSSTNSTNSTNSTNSLNGSOTSIGSTG--PWRSLNNKNSFLGTPFHRRKM 612  
 QY 572 QVPAEEMSSLTPE--SSPELAKSNWPGNF---ISLDEKEOIFVLKDKXPLSTIKADLVHA 627  
 Db 613 SNGTAEEDSDSDQIMIDTDLVKKSWFESLSSMSVREDTHCVPOCKTLNLSKAEILRA 672  
 QY 628 FLTSPLSLSHSVLSQTSFRAEYK--AASGGPVPFOKAPRFQVDDISSSGPEPPRRDG--SGG 684  
 Db 673 FLQIHLSHLSHVQONCFREYKRGKPTGVSQVSPRGIKMWDDI-----PSPQOVIAGE 726  
 QY 685 GGIVSYVTFLISGSPSRFRFVYETLOAOILSHDOPSYOALADENKGA--OTRP 736  
 Db 727 TPTVYVOFVLLAGPVPRFRKRLVHEHLSAIL-----ONSTQORADROQAALMWPR 775  
  
 RESULT 9  
 Q95T82 PRELIMINARY; PRT: 701 AA.  
 AC Q95T82;  
 DT 01-DEC-2001 (TREMBlrel, 19, Created)  
 DT 01-DEC-2001 (TREMBlrel, 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel, 20, Last annotation update)  
 DE GH1047p.  
 GN CG6114.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephyroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID:7227;  
 [1]  
 SEQUENCE FROM N.A.  
 RA STRAIN-BERKELEY.  
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Fattah D., Flise E., George R.,  
 RA Gonzalez M., Guanini H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nuno J., Pacle J., Paragas V., Park S., Phanenavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Gelman S.,  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY060288; AA125327.1;  
 DR FlyBase: FBmn0036544; CG6114.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR Pfam: PF00069; Pkinase.1.  
 DR ProDom: PD000001; Euk\_pkinase.1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR AMP-binding; Transferase.  
 SQ SEQUENCE 701 AA; 76973 MW; 87830879883B2425 CRC64;  
  
 Query Match 29.9%; Score 1228; DB 5; Length 701;  
 Best Local Similarity 45.2%; Pred. No. 3; 7e-73;  
 Matches 281; Conservative 86; Mismatches 142; Indels 112; Gaps 22.  
  
 QY 177 MAS10GDSLLTSCGSPHYACPEVTKGKRYGGRADMSGCVLFPALLVGLPDDDNL 236  
 Db 1 MAS10GASMLTSCGSPHYACPEVTKGKRYGGRADMSGCVLFPALLVGLPDDDNL 60  
 QY 237 ROLLEKVRKGVFHMHEFTPPDCQSLRGMIEVPEPKRLSLEQIQNHMY-LGKHEEP-- 293  
 Db 61 ROLLEKVRKGVFHMHEFTPPDCQSLRGMIEVNPDRRLTLAEINHHNPVLTAGGGELE 120  
 QY 294 -PCLLEPAGGRVAKRSLPNSGELDPDVLVSNASLGCFRDRRLRLRSEEDNEKMTYY 352  
 Db 121 LPMNEV-----VQHNHTPTATAVDPDVLNATCSLGCPEKREKLLQELLSSHNTEKYTF 175  
 QY 353 ILDRKERTYSCDED--LPKRD--VYPPRRKRVDSPLSRHGKRRPRKSKSEVLSIT 406  
 Db 176 ILDRKERTYSCDED--LPKRD--VYPPRRKRVDSPLSRHGKRRPRKSKSEVLSIT 406

```

OY 407 DAGGGSPV-PTRRALM-----AQSORS-----RSVSGASTGLSSPLS-----447
DB 228 ---SSGSPLEPRQAFAFNFRSYSTRNHQRSPPTVTSVSVSSSYHSPTRCNSSMSAQQO 284
OY 448 -----PRSPV-----PSFSPSP-----AGD-----463
DB 285 ANAISRPSPPAGTRHSTYGDNRSGHSSVSTPHSHSOKSIEGDVYVVRERIERDS 344
OY 464 --EARGGSPSTKOT-LPSRGPBGAGROPPPARSTPLDPPGSPRSSGGTPLASP 520
DB 345 LQERGGSPRDRDGCIPGSPGNSGSETASPSVHHRANSGP-----TIAISMHPD 399
OY 521 LHTPRASPTGP---GTPP--PPSPGGVGGAAMRSRLNSIRMSFSGSPRHRKMOVPT 575
DB 400 DSNVYVNPNGSPMNNSSPMPSPCPTPGGQMLTKTLNKSFLSPSPFRHRKMOV-S 458
OY 576 AEMSSLTTPSSPELAKRSWFGNFIISLDEEQIFLYLKDPLSLIKADIVHAFLSPLS 635
DB 459 ADHY-HITPSSSELTKRSMFGNLTENKETFTLYKGPATVKAHLIHAFLSMAELS 517
OY 636 HSYLSQTSFRAEYKASG-GSPVFOKPYRFOVDISSSEGPSPRRDSSGGGCTSYVFTL 694
DB 518 HSYVSPTSFVREYKRNCGVMEQRHYKFOVDISAI-----CKQGDIDMLFALFTL 570
OY 695 ISGSPRRKRVETIOQLS 715
DB 571 LSGNIRFRRIEHIQSVCS 591

```

## RESULT 10

OBSKS5 PRELIMINARY; PRT; 833 AA.

```

AC OBSKS5;
DT 01-JUN-2002 (TREMUREL. 21, Created)
DT 01-JUN-2002 (TREMUREL. 21, Last sequence update)
DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
DE Putative serine/threonine protein kinase.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AK4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
  Lehmann R., Baumgart C., Parra G., April J.F., Gulgo R., Kumpf K.,
  Tunga B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
  "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RT Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC115685; AAL92711.1; -.
DR KJ.
SQ SEQUENCE 833 AA; 93467 MW; 51B64036C90C0F94 CRC64;

```

Query Match 21.3%; Score 874; DB 5; Length 833;  
 Best Local Similarity 50.1%; Pred. No. 1.3e-49;  
 Matches 168; Conservative 67; Mismatches 93; Indels 7; Gaps 4;

```

OY 31 VGRVLEKTLGKGTGLVGVHCTGQVAKIVNREKI--SESYLMKVEREIAIKLI 88
DB 6 VGRVLEKTLGKGTGLVGVHCTGQVAKIVNREKI--SESYLMKVEREIAIKLI 65
OY 89 EHPHYVLEKTLGKGTGLVGVHCTGQVAKIVNREKI--SESYLMKVEREIAIKLI 148
DB 66 DHPHYVLEKTLGKGTGLVGVHCTGQVAKIVNREKI--SESYLMKVEREIAIKLI 125
OY 149 SYEICHRLKPEMLLDEKNNIRIADFGMSLOVSDLSLETSCGSPHYACPEVYKGEKYD 208
DB 126 SRNICHRLKPEMLLDEKNNIRIADFGMSLOVSDLSLETSCGSPHYACPEVYKGEKYD 185
OY 209 GRRADMSSCVLLFALLVGALEPDDNLRLQLEKVRGVFHPHFIPDDCSLRLGMIEV 268
DB 186 GOKADVMSCGVLLFALLVGALEPDDNLRLQLEKVRGVFHPHFIPDDCSLRLGMIEV 245
OY 269 EPERLSLEIQKHPWYLGKHEPDCLEPAFGRRVAMRLSPNGELDDVLESMAISLGC 328

```

```

DB 246 DPKRISIKIKHPWVFNQ---FOKATPVEEINAEPLVDYSQIDEDIFSIAMLG- 301
OY 329 FRDRERLHRELSEENQEKMTYLLDRKREYPS 363
DB 302 VGTIDEVKQOLVSNQKS-ATLITRYLLEERKFDPS 335

```

## RESULT 11

OBSKS5 PRELIMINARY; PRT; 1371 AA.

```

AC OBSKS5;
DT 01-NOV-1999 (TREMUREL. 12, Created)
DT 01-OCT-2001 (TREMUREL. 18, Last sequence update)
DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
DE KIAA0999 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirose M.,
  Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
  "Prediction of the coding sequences of unidentified human genes. XIII.
  The complete sequences of 100 new cDNA clones from brain which code
  for large proteins in vitro."
RT DNA Res. 6:63-70(1999).
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB023216; BAA76843.2; -.
DR HSBP; Q63450; IAO6.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Serine/threonine-protein kinase.
FT NON_TER
SQ SEQUENCE 1371 AA; 149525 MW; FOFBA385B8226158 CRC64;

```

Query Match 20.3%; Score 832; DB 4; Length 1371;  
 Best Local Similarity 28.6%; Pred. No. 1.5e-46;  
 Matches 246; Conservative 115; Mismatches 280; Indels 220; Gaps 25;

```

OY 2 SSGAKEGGGSPVHLPHP-----HHHPQOHQ-----YVGPRLEKT 39
DB 62 AAGAGTGGAGPAGRLPPRPGSPADPAVSPAAGPRPPAPSRGMPARIGYEIEDRT 121
OY 40 LGKQGTGLVGVHCTGQVAKIVNREKLSVLMKVEREIAIKLIHPHVLKLDV 99
DB 122 IGRNENAVKRAHLYKAKVAIKIIDTQDLENLKKIFREVOIMKMLCHPHIIRLYOV 181
OY 100 YENKRYLVLEHVSGLGELYLYKKGHLTPKEARKFFROLYVSLDPCHSYISICHRLKP 159
DB 182 METFRMITYLVLEHVSGLGELYLYKKGHLTPKEARKFFROLYVSLDPCHSYISICHRLKP 241
OY 160 ENLLDENNNIRIADFGMSLOVSDLSLETSCGSPHYACPEVYKGEKYDGRADMSSCGV 219
DB 242 ENLLDENNNIRIADFGMSLOVSDLSLETSCGSPHYACPEVYKGEKYDGRADMSSCGV 301
OY 220 ILFALLVGALEPDDNLRLQLEKVRGVFHPHFIPDDCSLRLGMIEVPEPKRSLDOI 279
DB 302 VLYVVGALPFDGSLTQNLRAVYLSGKFRIPFWMSTCEHLIRMLVLDPKRSLMEOI 361

```

```

QY 280 QKHPWLGKHEPDPCLPEAPGRVA-MRSLPSNGELDP---DYLSEMASLGCPRDEREL 335
DB 362 CKHKMKMG--DADPNND---RLNACQQLKEKROVDPLNEDVLLAMEMDGL--DKEQT 413
QY 336 HRELSEEMOEKMIYLLIDRKERYPCEDDLP--PR-----NDV 375
DB 414 LQSLSDAYDHSATYSLCDRKHRLTKLRLGALPSMPRALAFQAPVINOAEQAGTAMNI 473
QY 376 DPEPRKVDSPMLSRHGRKRRERKSM---VLST-TDAGGGGSPVPRRLAEMOHS----- 427
DB 474 SVFOVQLIN-----PENQIVPDGLNLDSDGEEPSPEALVRYLSMRHRYGVA 523
QY 428 -QRSNSVSGASTGLSSPLSPSPVFSPEPGAGDEAKNGGSPSTKQTLPSRGPRG 486
DB 524 DPTIEMEDQLKLPGFPGVNPOAPFLQVAP-----NVNFMHNLPRMQLDPT 571
QY 487 GAGEP-----PPSARSPTLPQPGSPRSRSGT--PLHSP--LHTRPA--SP--TGTPT 534
DB 572 GQLEYKESLQPTPLQLNGMPLGRASDGGANILQHAQQLKRRPSPPLVWTTPAV 631
QY 535 TPSPGCGVGAAMRSRLNSTRNSFLGSPRHRKMQVPTAEKMSLTPESPSELAKRS 594
DB 632 -----PAVTPVDESSDGPDDQEAQVSTY 656
QY 595 WGFNFISLDEKQFLVYLKDRPLSLKADIVHAFSLISLSVLSQTSFRAEYKASGP 654
DB 657 KDSNTLHLEPTER-----FSPVRRFSDDGASIAQFAKHLKMGNN 695
QY 655 S-----VFOKPVROVDISSSGPEPSPRRDGGGGIYVTPTLISGPRRRFR 704
DB 696 SSIKQLOECERQLOKMYGGQID-----ER 719
QY 705 VETIOQLSTHDPQSVQALADENGAQTPACAP-----RSLQ----- 745
DB 720 TLETKOQHMLYQEQHQLIQOQIIDSICPPSPPIQACENOPALITLQQLRIQP 779
QY 746 --PPGPRDPELSSPRKGP 764
DB 780 SSPPNHNPNNHLFRQPSNPP 800

```

RESULT 12

Q96RCO PRELIMINARY; PRT; 691 AA.

AC Q96RCO; 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, last sequence update)

DR 01-JUN-2002 (Tremblrel. 21, last annotation update)

DE Ser/Thr protein kinase PAK-1beta.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI\_TaxID=9606;

RP SEQUENCE FROM N.A.

RA Sun T.O., Lu B., Feng J.J., Reinhard C., Jan Y.N., Faull W.J., Williams L.T.,

RT "PAK-1 is a Dishevelled-associated kinase and a positive regulator of Wnt signaling."

RL Nat. Cell Biol. 3:0-0(2001).

DR EMBL; AF387638; AAK82368.1; -

DR InterPro: IPR000719; Euk\_kinase.

DR InterPro: IPR001772; Kinase\_Cterm.

DR InterPro: IPR002290; Ser\_thr\_kinase.

DR Pfam; PF00627; UBA domain.

DR Pfam; PF00627; UBA domain.

DR Pfam; PF00627; UBA domain.

DR ProDom: PD000001; Euk\_kinase.1.

DR ProSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN.1.

DR ProSITE; PS00101; PROTEIN\_KINASE\_DOM; 1.

DR ProSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN.1.

```

KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 691 AA; 77631 MW; 2216B4A7BC931BE CRC64;
Query Match 20.28; Score 828; DB 4; Length 691;
Best Local Similarity 30.94; Pred. No. 1,le-46;
Matches 232; Conservative 115; Mismatches 261; Indels 142; Gaps 24;
QY 30 YGPPRLKTLGKQOTGLKGVICINGQVAKIVREKLSSEVLMKVEREIALILIE 89
DB 16 HIGNYRLKLTIGNGNFKAVLARIHLTGKEVAKIIDKTLNSSLQKLEPRVIMKVLN 75
QY 90 HPPVAKIHDYENKYYLYLVLEVSGGELPDYLVKGRLLTPKARKFPQIVSALDCHS 149
DB 76 HPPVAKIHDYENKYYLYLVLEVSGGELPDYLVKGRLLTPKARKFPQIVSALDCHS 135
QY 150 YSICHDLEKENTLDEKNRIADFGMSLOYGDSILETSCGSPHYACEVYKGEYDG 209
DB 136 KEIVHNDIAKENTLDEKNRIADFGMSLOYGDSILETSCGSPHYACEVYKGEYDG 195
QY 210 RRADMMSCGVIIFALLVGALEPDDNMLKQLEKVKGVFHPFIPPDQSLRGMIYE 269
DB 196 PEYDVMSLVGVIYTLVSGSLPFDGQMLKELREVLNKGKRIPIRYMETDLENLKLFLIN 255
QY 270 PEKRLSLEQIOKHPWYLGKHEPDPCLPEAPGRVAMRSLPSNGELDPVLESMASTL 326
DB 256 PSKRGTLQIMDRMANNV--HEDDELKPYVEPLPDYK-----DPRTELVMGM 302
QY 327 GCFRDERELHRELSEEMOEKMIYLLIDRKERYPCEDDLPPLNDVDPKRYKVDSP 385
DB 303 G--YREETIDSLVQORV--EVMAIYLLLGYSSELBGTITLKRPSADLTNSSAPSPS 359
QY 386 -MLSRHGRKRRPKSMELVISTDAGGSGVPYTRALE---MAHSORSVSGASTGLS 441
DB 360 HKVQASVSNAPKQRF-----SDQAGPAIPTSNYSKKTQSNNAEKRPEDRESGRK 412
QY 442 SSPLSPSPSPVFSPEPGAGDEARGGSPSTKQTLPSRGPRGAGAEQPPPPASRSTP 501
DB 413 AS--STAKVPA--SPLPGL--ERKKTTPPTSTNVSLSTSN-----NSRNSP 453
QY 502 L--PGPPSPRSSSGGTPLASPLHTPRASPT-----GTPGT----- 535
DB 454 LLERASLQASIQNGKSTAPORVPVSPSAHNISSGAGDRTNPFPGVSSRSTFHAGQ 513
QY 536 -----PPSPGCGVGAAMRSRLNSTRNSFLGSPRHRKMQVPTAEKMSLTPESPSELAKRS 594
DB 514 LROYRQDQNLPGVTPASPSGHSGR--KQASGSITFSK--TSKFEVRNLNLPESKDRVE 569
QY 581 SLTP-----ESSPELAKRSW--FGNFTSLDEKQIFLYLKDPLSLIKAD 623
DB 570 TLRHVVGSGGNDKEKEEFBEAKPRSLFTWSMKTSTSMENEMMRILRYLDANSQSE 629
QY 624 IVHAFSLISLSVLSQTSFRAEYKASGSPVQKRYROVDISSGEPSPRRDGGSG 683
DB 630 LHREKYMML--CMGTGHEP-----VQWEMEVC-----LPRLSLNG 665
QY 684 GCGIYSTPTLISGSPRRFRVETIOQL 713
DB 666 -----VREKRISGTSMAFKNISKIANEL 689

```

RESULT 13

Q15524 PRELIMINARY; PRT; 745 AA.

AC Q15524; 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-FEB-1997 (Tremblrel. 02, last sequence update)

DR 01-JUN-2002 (Tremblrel. 21, last annotation update)

DE Serine/threonine protein kinase.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI\_TaxID=9606;

RP [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE-COLON; PubMed-9730619;  
 RX MEDLINE-98399868; Navaro E.;  
 RA Esplanosa L.; Navaro E.;  
 RT "Human serine/threonine protein kinase EMK1: genome structure and cDNA  
 cloning of isoforms produced by alternative splicing."  
 RT Cytogenet. Cell Genet. 81:278-278(1998).  
 RL CYTOGENET. CELL GENET. 81:278-278(1998).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: X97630; CAA6229.1; -.  
 DR HSSP: O63450; 1A06.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001772; Kinase\_Cterm.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR InterPro: IPR000449; UBA\_domain.  
 DR Pfam: PF02149; KAI; 1.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00627; UBA; 1.  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR SMART: SM00165; UBA; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 KW SEQUENCE 745 AA; 83204 MW; CEEA02EE762EC5DF CRC64;

Query Match 20.1%; Score 827; DB 4; Length 745;  
 Best Local Similarity 30.9%; Pred. No. 1.5e-46;  
 Matches 245; Conservative 114; Mismatches 259; Indels 176; Gaps 26;

30 YGPRLEKTLGKGTGLVHCITGOKAIAKIYREKISESVLMKVEREIALKLE 89  
 16 HIGNRLKTIKGNFAVKLAHILTGKEVAKIIDKTQNSSLOKLFREVRIMKVLN 75  
 90 HPVILKLDVYENKRYLLVLEHVSGBGLFDVLYKKRLTPREARKREFQIYALDFCS 149  
 76 HPNIVKLEFVETETKLYLMEYASGGEVDYLAHGRMKKEKARAFKQIYSAVQYCHQ 135  
 150 YSICHRDLKPEMLLDKNNIRIADFGMASLQVGSILFETSCGSPHYACPEVIGEKYDG 209  
 136 KTIYHRDLKAEMLLDADNMNITADFGFSNEFTFGKLDTPCGSPHYAPALFEGKKYDG 195  
 210 RRADMSCGVILFALLVGLPFDDNLRQLLEKVRGVHMFHPIPPDQSLRGMIYE 269  
 196 PEVDVMSGLVILTYLTVSSSLPFDGONLKELEKREVRLEKGRIPRYMSTDCENILKFKLLIN 255  
 270 PEKRLSEIOIKNHPYLGKKEHPD---PCLPAPGRVAMRSLPSNGELDPVLESMSL 326  
 256 PSKRGLLEQIMKDRMNVG--HEDELKPYVEPLPYK-----DPRTELAVSM 302  
 327 GCFRDERLHRELSEENOEMKIYLLDRKERYPSCEDODLPPRNDV----- 376  
 303 G--YTREIQLDSLVGQRYN--EVMATYLLLGKSSLEGGDTITLKPRPSADLTNSAPSPS 359  
 377 -----PPKRYDS-----PMLSRHK-----RPE-----RKM----- 400  
 360 HKVQSVSANPKQRKRSDDGAPAIPTSNYSKKTQSNNAENKREEDREGSKRASSAKY 419  
 401 -----EVLSTDAGGGSPVPTRALLEMA--OHSQRSVSG--A 436  
 420 PASPPLGLEKKKTTPTPTSTSVLSTINRSRNPFLERASLGQASIQNGKDSLMPGSR 479  
 437 STGLSSPLSSPR-----SPVSESPGAGDEARGGSEFTSKTQTLPSRGP----- 483  
 480 STASASAAVSAARPRQHKMSASVHFNKASGLPTESENCEVPRPSTAPQKVPVAPSAH 539  
 484 ----RGAGAGD-----PPPSASRPLPGPSPRSGGGLPLSLPHTPRASPTGPTTP 537  
 540 NISSGAPRTNFPKAVSSRSTHAGOLROYRQONLPY-----GVT-P 583  
 538 PSPGGVGAAMRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLRSLR 584  
 584 ASPGCHGGR--RQASGSIFSKF--TSKVFVRNINPEPSKDRVETLRPHVVGSGGNDK 639

QY 585 ----ESSPELAKRSM-FGNFISLDEKEDILYVLDKPLSLKAIIVHAFSLPSLSHVL 639  
 DB 640 EEFREAPRSLRFTWSMKTSTSMENEMMRIRYVLANDSCSELHEKXMLL--CMHGP 697  
 QY 640 SQTFRAYKASGSPVQKQVRFQVSDISSGEPSPRRSGGSGGIGYSVTFTLISGPS 699  
 DB 698 GHEDF-----VQMEVEYCK-----LPRLSLNG-----VRFKRIQSTS 729  
 QY 700 RRFKRYETIOAQL 713  
 DB 730 MAFNIAKINEL 743

## RESULT 14

096HB3 PRELIMINARY; PRT; 755 AA.  
 AC 096HB3  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Similar to EukL motif kinase.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PLACENTA;  
 RA Strusberg R.;  
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: BC008771; AA08771.1; -.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001772; Kinase\_Cterm.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR InterPro: IPR000449; UBA\_domain.  
 DR Pfam: PF02149; KAI; 1.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00627; UBA; 1.  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
 DR ATP-binding; Kinase; Transferase.  
 KW SEQUENCE 755 AA; 84332 MW; 493CD8C6DC3A4C06 CRC64;

Query Match 20.1%; Score 827; DB 4; Length 755;  
 Best Local Similarity 30.8%; Pred. No. 1.5e-46;  
 Matches 247; Conservative 112; Mismatches 261; Indels 182; Gaps 26;

30 YGPRLEKTLGKGTGLVHCITGOKAIAKIYREKISESVLMKVEREIALKLE 89  
 16 HIGNRLKTIKGNFAVKLAHILTGKEVAKIIDKTQNSSLOKLFREVRIMKVLN 75  
 90 HPVILKLDVYENKRYLLVLEHVSGBGLFDVLYKKRLTPREARKREFQIYALDFCS 149  
 76 HPNIVKLEFVETETKLYLMEYASGGEVDYLAHGRMKKEKARAFKQIYSAVQYCHQ 135  
 150 YSICHRDLKPEMLLDKNNIRIADFGMASLQVGSILFETSCGSPHYACPEVIGEKYDG 209  
 136 KTIYHRDLKAEMLLDADNMNITADFGFSNEFTFGKLDTPCGSPHYAPALFEGKKYDG 195  
 210 RRADMSCGVILFALLVGLPFDDNLRQLLEKVRGVHMFHPIPPDQSLRGMIYE 269  
 196 PEVDVMSGLVILTYLTVSSSLPFDGONLKELEKREVRLEKGRIPRYMSTDCENILKFKLLIN 255  
 270 PEKRLSEIOIKNHPYLGKKEHPD---PCLPAPGRVAMRSLPSNGELDPVLESMSL 326  
 256 PSKRGLLEQIMKDRMNVG--HEDELKPYVEPLPYK-----DPRTELAVSM 302  
 327 GCFRDERLHRELSEENOEMKIYLLDRKERYPSCEDODLPPRNDV----- 376  
 303 G--YTREIQLDSLVGQRYN--EVMATYLLLGKSSLEGGDTITLKPRPSADLTNSAPSPS 359

```

QY 377 -----PERRVDS-----PMLSRHGK-----RPE-----RKSM-----400
Db 360 HKVORSVSNPKNRFRFDQAAPAIPTNSNSYSKKTOSNNAKREERESGRKASSTAK 419
QY 401 -----EVLSTIDAGGGGSPVPRRLALEMA--OHSORSRSVSG--435
Db 420 VPASPLPGLERKTTTPPTSTNSVLTSTNSRNSPLERKSLGQASLONGKDSLMPGSR 479
QY 436 ASTGLSSPLSPR-----SPVSEFEPGAGDEARGGSPSTKOTLPISRG-483
Db 480 ASTASASAAVSARPRQHOHKSMSASVHPNKAAGLPRESNCEVPRPTAPRVPVASSA 539
QY 484 -----RGGGAGEO---PPPSARSTPLPGPPSGSPSSGCTPLHSPHTPRASPTGTPTT 536
Db 540 HNISSGGAPDRTNPRGVSSRSTFHAGQLQVDDOONLRY-----CVT-583
QY 537 PPSPGGVGAAMRSRLNSTRNSFLG-----SPRHRKMOVPTAE--MSLTP-----584
Db 584 PASPGHSGR--RGASGSIKFTSKFYARNLSFRARNLNEPESKDRVETLRPHVVG 641
QY 585 -----ESSPELAKRSW-FGNFISLDKEQIFLVLDKPLSSIKADIYHAFISI 631
Db 642 SGGNDKEKEEFREAKFRSLFTWSMKTSSMEPNEMAREIRKVLANSQSELHEKTM 701
QY 632 PLSHSHVLSQTSFRAEYKASGSPVQKPVROVDISSSGPEPPSRDGGGGGIYSVT 691
Db 702 ---CMHGTFGHEP-----VQWEMEVC-----LPRLSLNG-----VR 731
QY 692 FTLSHSPRRERKRVETIOAL 713
Db 732 FKRISTGMAFKINASKIANEL 753

RESULT 15
Q9BYD8 PRELIMINARY; PRT; 688 AA.
AC Q9BYD8;
DT 01-JUN-2001 (Tremblrel, 17, Created)
DT 01-JUN-2001 (Tremblrel, 17, last sequence update)
DE MAP/microtubule affinity-regulating kinase like 1.
OS MARKL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=21226021; PubMed=11326310;
RA Kato T., Satoh S., Okabe H., Kitahara O., Ono K., Kihara C.,
RA Tanaka T., Tsunoda T., Yamaoka Y., Nakamura Y., Furukawa Y.,
RA Isolation of a novel human gene, MARKL, homologous to MARK3 and its
RA involvement in hepatocellular carcinogenesis.
RT Neoplasia 3:4-9(2001)
CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049127; BAB39380.1;
DR HSSP; O63450; 1A06.
DR InterPro; IPR000719; Euk_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF00627; UBA; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYRKc; 1.
DR SMART; SM00165; UBA; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

```

```

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 688 AA; 75261 MW; A03B5A7943ACD086 CRC64;
Query Match 20.1%; Score 826; DB 4; Length 688;
Best Local Similarity 32.4%; Pred. No. 1,5e-46;
Matches 235; Conservative 93; Mismatches 223; Indels 174; Gaps 20;
QY 25 PQHAQVGPYRLKTLKGLGQTLVGLVHCITGQVAKIKIVNEKLSSEYLMKVEREIAI 84
Db 50 PEEQAPVGVNRLRTITGKNSAKVLAHILTGREVAIKIIDTQNLPPSLQKLFREYRI 109
QY 85 LKIEHPHYLKLHDYENKRYLYLYLEHNSGGELFDYLVKGRLLPREAKRFQIYSAL 144
Db 110 MGLNHPNIVKLFVETETKTLVMEYASAGEVFYLVSHGMEKEKARAKRQIYSAV 169
QY 145 DECHSYSLCHRDLPENLLDEKNIRIDFGMASIQVGSLSLETGSGSPHYACPEYIK 204
Db 170 HCHQKNIVHRDLKKNENLLDAENIKIDFGSNEFTLGSKLDTGSPPYVAPELFG 229
QY 205 EKYDGRADMGCGYILPALLVGLPDDDNILQLLEKYKGVFNHPHPIPPDCSLRG 264
Db 230 KYDPEVDIWSLGVILTVLVSGLPFDGHNILKEIRVLRGVYVPYVSTDCESILR 289
QY 265 MIEVEPERLSLEQIOKHPWYLGKHEPDCLPEAPGRVYAKSLPSNDELDPVLESMA 324
Db 290 FLVLPNAPKRCITLQIMKOKWITGYE-----GELKPYTEDEEDFGTKRIEVAV 339
QY 325 SLGCRDRERLHRELSEENQEMITYLLDRKERYPSCEDQDLPFRNDVDPKRYVDS 384
Db 340 GMG--YTREEKESLTSOKYN-EVATYLLGKRTKE---EGGD---RARGLALARVVA 389
QY 385 PMLSRHGKRRPERKSMELVSTIDAGGGSPVPTRALALEMAHOSRSYSGASTGLSSP 444
Db 390 PSDTTNG-----TSSKGTSHSKGQSSSTYRORRHS---DECGPPAP 432
QY 445 LSSPRPVSEFSPGAGDEARGGSPSTKOTLPSSRGPGGAGEOPPPPARSTPLPG 504
Db 433 LHPKRSPTST-----GEALKEERLPGR-----KASCTAGSG 465
QY 505 PPGSPRSGGCTPLHSPHTP-----RASPTGTPTTTPPPSGGVGAAMRSRLNSI 556
Db 466 SRGLPSS---PMVSAHNPKAETPERKRDSTPTNNLP-----SMATR 508
QY 557 RNSFLGSPR-FHRRKMOVPAEEMSSILPESSPELAKRSWFGNFIISLDKEQIFLVLDK 615
Db 509 RNTYVCTERPGAERPSLLPNKENSSTGTPRYP-----541
QY 616 PLSIKADIYHAFIPLSTLSHVSLSQTSFRAEYKASGSPVQKPVROVDISSSEGP 675
Db 542 -----ASPS-SHSLAPPSGERS--RLARGSTI-----RSIFHGQV 574
QY 676 SPRRDGGGGGIYSVFTLISGPRFRKRVETIOALLSHDPSVOALADENGAQTR 735
Db 575 RDRRAGGGGG-----GYONGP-----PASPTLAHE---AAPL 604
QY 736 PAGAP 740
Db 605 PAGRP 609

```

Search completed: April 16, 2003, 12:32:07  
Job time : 104 secs

GenCore version 5.1.4-P5-A578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 12:22:22 ; Search time 28 Seconds  
(without alignments)  
1152,449 Million cell updates/sec

Title: US-10-003-690-2

Perfect score: 4108  
Sequence: 1 MSSGAKEGGGSPVHPLP.....PRGPPKDKKLATNGTLP 778

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SWISSPROT\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Query length	DB ID	Description
1	836	20.4	1518	1	KKK1_YEAST
2	816	19.9	786	1	SNIL_HUMAN
3	797	19.4	713	1	RP78_HUMAN
4	795.5	19.4	512	1	K110_ARATH
5	788	19.2	774	1	KEMK_MOUSE
6	784	19.1	776	1	SNIL_MOUSE
7	770.5	18.8	779	1	SNIL_MOUSE
8	760.5	18.5	1142	1	GIN4_YEAST
9	757.5	18.4	576	1	SNF1_SCHPO
10	753.5	18.3	633	1	SNF1_YEAST
11	730	17.8	619	1	SNF1_CANTR
12	725	17.6	550	1	AAK1_HUMAN
13	722	17.6	548	1	AAK1_MOUSE
14	722	17.6	915	1	KCC4_YEAST
15	720	17.5	611	1	SNF1_MOUSE
16	718	17.5	552	1	AAK2_HUMAN
17	715	17.4	620	1	AAK2_MOUSE
18	710.5	17.3	520	1	SNF1_CANTL
19	672	16.4	891	1	KIN1_SCHPO
20	666.5	16.2	502	1	RR11_SECE
21	636.5	15.5	622	1	YNA3_CAEEL
22	632.5	15.4	661	1	CDRL_SCHPO
23	617.5	15.0	593	1	CDRL_MOUSE
24	594.5	14.5	1064	1	KIN4_YEAST
25	570.5	13.9	714	1	HUNK_HUMAN
26	569	13.9	800	1	KIN4_YEAST
27	556.5	13.5	1147	1	HUNK_MOUSE
28	555.5	13.5	714	1	HUNK_MOUSE
29	550	13.4	664	1	KCCB_HUMAN
30	521	12.7	1050	1	ULK1_HUMAN
31	518.5	12.6	353	1	ASR2_ARATH
32	497	12.1	499	1	KCCD_HUMAN
33	495	12.0	533	1	KCCD_MOUSE

34	482.5	11.7	370	1	KCC1_HUMAN	Q14012 homo sapien
35	482.5	11.7	1051	1	ULK1_MOUSE	Q70405 mus musculu
36	479.5	11.7	542	1	KCCB_MOUSE	P28652 mus musculu
37	479.5	11.7	542	1	KCCB_MOUSE	P08413 rattus norv
38	479	11.7	295	1	KMIC_DICDI	P25323 dicystosell
39	478.5	11.6	374	1	KCC1_MOUSE	Q63450 rattus norv
40	478.5	11.6	478	1	KCCB_HUMAN	Q96457 homo sapien
41	478.5	11.6	478	1	KCCB_MOUSE	P1275 rattus norv
42	477.5	11.6	472	1	KCCB_HUMAN	Q1355 homo sapien
43	477	11.6	363	1	ASR1_ARATH	P43291 arabidopsi
44	474.5	11.6	406	1	KPCB_MOUSE	Q9db30 mus musculu
45	471.5	11.5	527	1	KCCG_MOUSE	P1730 rattus norv

## ALIGNMENTS

RESULT 1  
ID KKK1\_YEAST STANDARD: PRT: 1518 AA.  
AC P34244;  
DE 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DF 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Probable serine/threonine-protein kinase YKL101W (EC 2.7.1.-).  
GN YKL101W OR YKL453.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_Taxid:4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN: S288C;  
RX MEDLINE:9407677; PubMed:8256524;  
RA Pallier C., Valens M., Puzos V., Fukuhara H., Cheret G., Sor F.,  
RA Boloitin-Fukuhara M.;  
RT "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI  
RT physically localizes the MRB1 gene and reveals eight new open reading  
RT frames, including a homologue of the KIN1/KIN2 and SNL1 protein  
RT kinases."  
RT Yeast 9:1149-1155(1993).  
RT -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC NIM1 SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC EMBL: X71133; CA550456.1;  
CC EMBL: Z28101; CA61941.1;  
CC PIR: S37928; S37928.  
CC PIR: S39084; S39084.  
CC HSP: O63450; 1A06.  
CC SGD: S0001584; YKL101W.  
CC InterPro: IPR000719; Euk\_pkinase.  
CC InterPro: IPR002290; Ser\_thr\_pkinase.  
CC Pfam: PF00069; pkinase; 1.  
CC ProDom: PD000001; Euk\_pkinase; 1.  
CC SMART: SM00220; S\_TKC; 1.  
CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
CC PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
CC Hypothetical protein; transferase; Serine/threonine-protein kinase;  
KW ATP-binding.  
KW DOMAIN 81 369 PROTEIN KINASE.  
FT ND BIND 87 95 ATP (BY SIMILARITY).  
FT BINDING 110 110 ATP (BY SIMILARITY).  
FT ACT\_SITE 239 239 BY SIMILARITY.  
SQ SEQUENCE 1518 AA: 169592 MM: 803F84FF731241DD CRC64;

Query Match 20.4%; Score 836; DB 1; Length 1518;  
 Best Local Similarity 28.5%; Pred. No. 1.8e-26;  
 Matches 256; Conservative 135; Mismatches 261; Indels 246; Gaps 29;

31 VGPYRLKTLGGQGLVLTGVCITGOKAIVNREKL----- 70  
 78 VGPWIKTLGGKSSGVRVLAQKMTGOLAKIVPKKAFVHCNSNGTVPNSYSSSVMT 137  
 71 -----SESVLMK-----VEREIALKLEIHPVTKLDHYVKNKTYLYLEHVS 114  
 138 SNVSSPSISRSNHSQNPNGIEREIVIMKISHNVALEEVWKNKSELYLVLEVD 197  
 115 GSELDTLYKKRRLPKPKRFFROJVSALDFCHYSICHRLDKPENILADEKN-NIRIA 173  
 198 GSELDTLYVSKRLPEREAIHYEKOIVECVSYCHSFNICHRLDKPENILDKKRRIRKIA 257  
 174 DFGMASIQVDSLETSCSPHYACPEVIGKEGYDGRADMSGCVTLFALLVGLPDD 233  
 258 DFGMALELPNKLTKSCSPHYASEIWMGRYPHGGPSDWSGCVLFLGLHLPFND 317  
 234 DNLKOLEKVRGVFHPHPIPPDCQSLRGMIIEVERKRLSEIOIKHPW----- 284  
 318 DNKKTLTKVSGKYOQMSNLSEARDLTKLIVDEPKRITTOELKHLKHYDLPV 377  
 285 -----YLGKHEPDPCELEPAGRRVAMRSLPNSGELDPVLESMAISGCFRDR 333  
 378 NKVLKMKMKMMARKKSNDS--LHLLNVSFIVTLHSGEIDESTLRSLQILMHGVSRE 435  
 334 RLHRELNSEENQEKMIYLLDRKERY-----PSCEDQDPPRNDVDPKRKVDSPMS 388  
 436 LITAKLLKPKMSEKLFYLLQYKORHSISLSSSENKSKTESVNPRIEYASKTAN 495  
 389 RHGRPERKMEVLSTIDAGGSGPVPTRALEMAQHSQ-----RSSVSGASTGL-- 440  
 496 NTGLR-----SENNDVTKLSLET--HSEDTSTVQNNAIIGVNTENA 537  
 441 -----SSPLSPRS-----PVFSFPEPGADDEAGGSPSTKQTF 477  
 538 PVLAQKQSFINTLSOPESDKAEAVTLPAIPAFNMS-----SKIFRNSTSISS 590  
 478 LPSGPRGGGGEPPPSANST-----PLPGPPGSP-----RSSGGRP-----LHSP 520  
 591 KRSRSLSLNSRLSLASSTSHREYHDMEMPLPOLPKSPRSYLSRAIHAASPTSIHS 650  
 521 LHTPRASPTGTPTPPSPGCGVGAAMRSLMSI--ANSLGSRFRFRKKQV----- 573  
 651 LSKRNIAAT-----VAKRKLQNSAKRSLSYLSQISIKRSLNLDLVF 694  
 574 -----PRAEKSSILTPRESSPE-----LAKRWFNGFIS--LDKEQILFVLKDK-- 615  
 695 DDELPSKPKPASEVWVKNKEPSHLESDDFELCDQILFGNALDRILEED-----NEKER 749  
 616 -----PLSSIKADIYHAFILSTLSHVSLSQTSFRAEYASGSPS--- 655  
 750 DPGORONDTKSSADFTTISGVSTNKNEGPEYP---TKLEKQFNMSYKPESEMGSL 805  
 656 ---VEQKPYARFOYDSSSEPEPSRRDGGGGGIVTFTLLISGSRFRKRYVETIO-A 711  
 806 SEPIFEK-----EWTLSSYLEQKPKR-----AALSDITNSFNKMNKQEKRIETKORE 856  
 712 OLLSTHDOFSVQALADEKNCQATRPAGAPRSIQP-----PPGAPPELSSPPR 761  
 857 QLOKKNDRPS-----PLKPIQHOELRVNSLPPNDGKPSLSLDPDR 896

DE Probable serine/threonine protein kinase SNFLK (EC 2.7.1.-).  
 GN SNFLK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 CX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-20289799; PubMed-10830953;  
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,  
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,  
 RA Ohts M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,  
 RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,  
 RA Reichwald K., Rump J., Schillabel M., Schudy A., Zimmermann W.,  
 RA Shintani A., Saeki T., Shibuya K., Kawasaki K., Asakawa S.,  
 RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,  
 RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,  
 RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,  
 RA Lehmann S., Borzjyn K., Gardiner K., Nizetic D., Francis F.,  
 RA "The DNA sequence of human chromosome 21.";  
 RL Nature 405:311-319(2000).  
 CC - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC SNF1 SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: AP001751; BA95536.1; -  
 DR HSSP; P24941; 1A01.  
 DR GeneW; HGNC:11142; SNFLK.  
 DR MIM; 605705; -  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR004449; UBA\_domain.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00165; UBA; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR Transferase; Serine/threonine-protein kinase; ATP-binding.  
 FT DOMAIN 27 281  
 FT NP\_BIND 33 41  
 FT BINDING 56 56  
 FT ACT\_SITE 149 149  
 FT ACT\_SITE 149 149  
 FT BY\_SIMILARITY.  
 FT BY\_SIMILARITY.  
 SEQUENCE 786 AA; 85252 MW; FB44EEFCAR87C9A CRC64;

Query Match 19.9%; Score 816; DB 1; Length 786;  
 Best Local Similarity 30.9%; Pred. No. 6e-26;  
 Matches 248; Conservative 103; Mismatches 301; Indels 150; Gaps 26;

31 VGPYRLKTLGGQGLVLTGVCITGOKAIVNREKLSESVLMKVEREIALKLEIHPVTKLDHYVKNKTYLYLEHVS 90  
 24 VGPYDLERTLGGNRAVAVKLARHRTKTOVALKIDKTRLDSSNLEKIVREYQILMKLNLH 83  
 91 PVHKLHDYENKRYLYLVLEHVSSELDPVYKGRGLPKPEARKEFRROJVSALDFCHYS 150  
 84 PHILKLYQVETKQMLYVTEPAKNGEMFDYLSNCHLSBNKARKFWQLSVEVCHDR 143  
 151 SICHRDLKPEINLLDEKNIRIA---DFGMASIQVDSLETSCSPHYACPEVIGKEGY 207  
 144 HIVHDLKTEINLLDGNMDIKLAGTEDFGNGFYKSGEPLSTWCGSPYAAPAEVFGKEKY 203  
 208 DGRADMSGCVTLFALLVGLPDDDNRLKLEKVRGVFHPHPIPPDCQSLGKMT 267

DB 204 EGPQDLSLVLYLVYVCGSLPFDGPNLPTLQRYLEGRFIPFMSQDCESLIRMLV 263  
 QY 268 VEPERKLSLEQLOKHPWYLGKHEPPDCELEPAG-RRVAMRSLPSN-GEIDDPVLESMA 325  
 DB 264 VDPARITIAQIQHMM-----RAEPLC-PGAPAFSAHSYTNLGDYDEQALGIMOT 317  
 QY 326 LGCFRDERLHRELSEENQEMKITYLLDRKERNVSCEDDLPRLNDVDPKRYDPS 385  
 DB 318 LGV--DROFTVESLONSSYNHFAIYLLERLEKYNNAO----- 355  
 QY 386 MLRRHG-KRRPERKSMVLSITDAGGSGVPTRALEMAOHRSRNSVSGASTG-LSS 442  
 DB 356 -CARPAPKPPRRSSLSLEVPQBELSDPFRPALLCPQPTQVSVLAQEMDCLOS 414  
 QY 443 S-----PLSPRRSPVFSFSP-----EPGADDERGSGSPSKQYTLRSPRGGCA 488  
 DB 415 SLQMPLEFPYDASGSGYFRPRVPSPSSILDTALISEARAQGLLEEDDTORSL-PSSTGR 473  
 QY 489 GEQPPRSASTPLRGP-----PGSPRSSGGTPLL----- 518  
 DB 474 RHTLAEVSTRSLPLTADCKYVSPSTTASPAAGTSSDCLTFSAKSPAGLSGTATQGLL 533  
 QY 519 ---SPLHTPRASP-TGTPGTTPPSPGGSGVGAAMRSLNSIRNSFLSGPPFHRKQV 574  
 DB 534 GACSPVRL--ASPLGSGSATPVLAQGLGAV-----LLPVSF-----QEGRRASDTS 581  
 QY 575 TAEEMSSITPSSPELAKRWFQNFISIDKREOFIVLKDPLSGIK---ADIVAFSLI 631  
 DB 582 LTQGLKAF---ROOLKRTTPTKGLGINKTKIGLAKOVCOAPASASGGLSPLAPAS 637  
 QY 632 P-----SLHSVLSQTS-FRAEYKASGSPVFOKPVRYOVDISSSEGPSP 677  
 DB 638 PGLGGAAGSREGNSILEVELEQRLQLDHNRAAP-----GCSQAPAPAP 684  
 QY 678 RR-----DSGGGGSIVTPT-----LISGSRKRYVETIQADLLSHDPSVALAD 727  
 DB 685 APFVAPDGGGAAPLPSTLLTSLGRLPRLLOTGASPVASAOULLDTH-----LH 736  
 QY 728 EKNGAOTRPAAGAPRSLOPPG 749  
 DB 737 IGTGPTALPAPVPRILARLAPG 758

RESULT 3  
 KP78 HUMAN STANDARD: PRT: 713 AA.  
 AC P27448:  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE putative serine/threonine-protein kinase P78 (EC 2.7.1.-).  
 GN P78.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Maheshwari K.K., Som S., Parsa I.;  
 RT Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.  
 CC -1- MISCELLANEOUS: MARKER PROTEIN LOST IN CHEMICALLY INDUCED  
 CC TRANSPLANTABLE CARCINOMA AND PRIMARY CARCINOMA OF HUMAN PANCREAS.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC NIM1 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: M80359; AAA59991.1; -  
 DR PIR: S27966; S27966.  
 DR HSP: 063450; 1A06.  
 DR Genew: HGNC: 6897; MARK3.  
 DR MIM: 602678; -  
 DR Interpro: IPR000719; Euk\_pkinase.  
 DR Interpro: IPR001772; Kinase\_Cterm.  
 DR Interpro: IPR002290; Ser\_thr\_pkinase.  
 DR Interpro: IPR000449; UBA\_domain.  
 DR Pfam: PR00069; pkinase.1.  
 DR Pfam: PR00067; UBA.1.  
 DR Pfam: PF02149; KAL.1.  
 DR ProDom: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00220; S\_LTC; 1.  
 DR SMART: SM00165; UBA.1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP.1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST.1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM.1.  
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding.  
 KW DOMAIN 56 307 PROTEIN KINASE.  
 FT NP\_BIND 62 70 ATP (BY SIMILARITY).  
 FT BINDING 85 85 ATP (BY SIMILARITY).  
 FT ACT\_SITE 178 178 BY SIMILARITY.  
 SQ SEQUENCE 713 AA; 79903 MW; 1CA78EB22620A228 CRC64;

Query Match 19.4%; Score 797; DB 1; Length 713;  
 Best Local Similarity 32.0%; Pred. No. 3; le-25;  
 Matches 237; Conservative 112; Mismatches 254; Indels 138; Gaps 28;

QY 30 YVGPRLKLETKGGTGKLVKLVCHCTGQKVAIKIYVREKLESYLMKVEREIAIKLE 89  
 DB 52 HIGNRRLKTKGKFNFAKAKLARIHLTGREVAIKIDKQLNPTSLQKLFREVIKILN 111  
 QY 90 HPHVLKLDVYENKKYLVLYLVHVSQGLPDIYLYKGRITPKAKFFQIYASALDFCS 149  
 DB 112 HPIYVLFVEVIEQKLYLIMEYASGKVFYLVAGRKKEKARKFQIYASAVQICQ 171  
 QY 150 YSLCHDLKPEMLILDEKNNIRADFMAISLVGDSILETSCSPHYACPEVIKGEYDG 209  
 DB 172 KRIYHDLKAEMLLDADMANIKIADFGSFNEFTVGGKLDTFGCSPPYAAPELFGKKYDG 231  
 QY 210 RRADWMSGVITFALLVGLPFDNNILROLLEKRYKGVFHPFIPPCOSLLRGATIEVE 269  
 DB 232 PEVDVMSLVGIYITLVSGSLPFDGQMLKELEKRYKIPYMYSTDEMLKRLVLN 291  
 QY 270 PEKRLSLEQLOKHPWYLGKHEPD---PCLPAPGRVARSLSPSNGEL---DPVLES 323  
 DB 292 PIRKGLTQIMKRWLNAG-HEEDLKPVEP-----ELDISQKRIIDIM 335  
 QY 324 ASLGEFRDERLHRELSEENQEMKITY-----YLLDRKERYPSCEDDLPRLND 374  
 DB 336 VGMG-----YSQELIGESLSKMKYDELTATYLLGRK-----SSEVRPSSD 376  
 QY 375 VD-----PPKRYDSPMLSHRGRPRPKSMVLSITDAGGSGSP---VPTRALMAQ 425  
 DB 377 LNNSTGSPHHKQVRSVSSQOKOR-----YSDHAGPGIPSVAYAPKRSQTSTAD 426  
 QY 426 HSOR-----SRVSAGASTGLSSPSLSPSPVFSFSEPGAGDEARGGSPSTKQTLPS 480  
 DB 427 SDLKEDISRKRSKSGAVGKG---IAPASMLGNASPNKADLPE-----RKSSVPS 478  
 QY 481 RGPGRGAGDEPPPSARSPTLPPEPPSPSSGGTPLHSLPHTRASPTGTPGTPPSP 540  
 DB 479 SNTASGCM-----TRNTYV-----SERTTDDR--HSVIONKEKENT-IPDQRTPVAS 524  
 QY 541 GGVGGAAMRSL-----NSIRNSFLGSPRRRRKQVPTAEKSSSLPSESPELAKRS- 594  
 DB 525 THSISSATPRKIPRKPTASRSTFHGPR-ERTATYNGPPASPSLSHEATPSQTSR 583  
 QY 595 ---WFGNFTS-LDKREQIFVLAD---KPLS-----SLK-----ADIVAFSLIP 632  
 DB 584 GSTTLFSKLTSLKTRSNVSAKODENKARPSLFTVSMKTTSSMDPGDMREIRKVL 643



Db 309 NHLIESLRNRRNGDTVTYIYLLDNRFRASGGYLAEFOETMEGTPRMHAEVASPVSH 368  
 QY 389 R 389  
 Db 369 R 369

RESULT 5  
 KEMK\_MOUSE STANDARD: PRT: 774 AA.  
 ID KEMK\_MOUSE Q05512;  
 AC 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Putative serine/threonine-protein kinase EMK (EC 2.7.1.-).  
 GN EMK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=93364122; PubMed=8358177;  
 RA Inglis J.D., Lee M., Hill R.E.;  
 RT "Emk, a protein kinase with homologs in yeast maps to mouse  
 chromosome 19.";  
 RL Mamm. Genome 4:401-403(1993).  
 CC -1. SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC NIM1 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X70764; CAA50040.1; -  
 CC DR HSSP: 063450; 1A06.  
 CC DR MGD: MGI:99638; Emk.  
 CC DR InterPro: IPR000719; Euk\_pkinase.  
 CC DR InterPro: IPR001772; Kinase\_Cterm.  
 CC DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 CC DR Pfam: PF00069; pkinase.1.  
 CC DR Pfam: PF02149; KAI.1.  
 CC DR Pfam: PD000001; Euk\_pkinase.1.  
 CC DR SMART: SM00220; S\_TKc.1.  
 CC DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP.1.  
 CC DR PROSITE: PS00108; PROTEIN\_KINASE\_ST.1.  
 CC DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM.1.  
 CC KM Transferrase; Serine/threonine-protein kinase; ATP-binding.  
 CC FT DOMAIN 53 304 PROTEIN\_KINASE.  
 CC FT NP\_BIND 59 67 ATP (BY SIMILARITY).  
 CC FT BINDING 82 82 ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 175 175 BY SIMILARITY.  
 CC SQ SEQUENCE 774 AA; 85874 MW; 02BF8D7BF443483A CRC64;

Query March 19.2%; Score 788; DB 1; Length 774;  
 Best Local Similarity 30.1%; Pred. No. 7.5e-25;  
 Matches 238; Conservative 119; Mismatches 261; Indels 172; Gaps 27;

QY 30 YVGPRLKTLKGGTGLVKGHCITGOKVAIKIVNREKLSSESLMKYERIALIKLIE 89  
 Db 49 HIGNRRLKLTICKGNFAKYKILARHILITGKVAVKIIDKTOQLSSLOKLFREVRIMKVN 108  
 QY 90 HHNVAKLDVYNNKKYIVLVEHVGSGLEFDLYVKKGRLTPEAKRFEPQIVSALDFCHS 149  
 Db 109 HNVIVKLEEVITEKTLVLMVEYASGGEFDLVAGNRKKEAKRAKQIYLVHYQYCHD 168  
 QY 150 YSICHRDLKPEMLIDDEKNNIRIADFGMASLQVGSLLFTSCGSPHYACPEVIGKEKYOG 209

Db 169 KPIVHRDLKAENILLDADANIKIADFGFSNEFTFGNKLDTFFCCSPYAAPELFQGGKIDG 228  
 QY 210 RRADWMSGCVITLFPALLVGLPFDDDNILROLLEKVKGVFMHFTIPDDCOSLIRGMIEVE 269  
 Db 229 PEYDVMSLGYITLVSGSLPFGQNLKEIREYRLRGKRIPIFYMSIDCENLLKFKLILN 288  
 QY 270 PEKRLSLEQIQKHPVTLGGHEPDPCLPEAPGRVAMRSLPSNGEIDPDVLESASIGCF 329  
 Db 289 PSKRGTLEQIMKDRMNVG--HEDD-----ELKPYV-EPLITIGP- 325  
 QY 330 RDR-----EKLHRELSEENOEKMIYLLDKRERYPCGEDDLPDRDVD---- 376  
 Db 326 RDVDVNGVNGHTEIDSLVGQRYN-EVMATYLLGLYKSSPEEDDTTLKPRPADLTNS 384  
 QY 377 -----PPKRYDS-----PMISRCK-----RRERKSMELVSI 405  
 Db 385 SAPSPSHKVOVSANPKORRSSDPAVPAIPTNSYSKKTQSNNAENKRDEEETGRASS 444  
 QY 406 TDAGGGGSPYP--TRALMAHQSRSVSGASTGLS--SSPLSSPSPVFSFSPBG-- 460  
 Db 445 T-AKVPAAPLPGIDRKKTTPA--PTNSVLISTNSRNSPLLD-RASLGQASIQNGKD 499  
 QY 461 ----AGDEANGGSPSTKTOTLPGRPGGAGBOP--PPSANSTPLPGPGSPBS 511  
 Db 500 SLTMGSRASASAAVSAARPHQHKMSGASVHPMKASGLPPTESNCEVPRSTAPOR 559  
 QY 512 SGTPLHSPHTPPASPPTGTGTT-----PPPSG 541  
 Db 560 --VPVAPSPAHNIISSSGADRTNPPRGVSSRSTFHAGQLRQVDOQNLPGYVTPASP 616  
 QY 542 GGVGAAMRSRLNIRNSFLSPFHRKMQVPAEE-MSLTP----- 584  
 Db 617 GHSGR--RGASGSITFSF--TSKFEVRNINLNEPSKDKVEITLRPHVSGGTDKDEEER 672  
 QY 585 ESSPELAKRSV-FGNFISLDEQITFLVLRKPKLSIADIVHAFSLPSLSHVSLSGT 643  
 Db 673 EAKPRSLFTWMSMTTSSMEPEMREIRKYLDANSCOSELHERYMLL--CVHGTPGHEN 730  
 QY 644 FRAEYKAGGGSVPQKPVRFQVDDISSGPEPSPPDSCGGGIGIYVFTLISGSPRERF 703  
 Db 731 F-----VQMEKEVCK-----LPLSLNG-----VAFKRISGSMAFK 762

QY 704 RVEVETQAOQL 713  
 Db 763 NIASKIANEL 772

RESULT 6  
 SNIL\_RAT STANDARD: PRT: 776 AA.  
 ID SNIL\_RAT O9R105; O9R081;  
 AC 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Putative serine/threonine protein kinase SNFilk (EC 2.7.1.-) (Salt-  
 DE inducible protein kinase) (Protein kinase KID2).  
 GN SNFilk OR SIX OR KID2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Adrenal gland;  
 RX MEDLINE=99330184; PubMed=10403390;  
 RA Wang Z., Takemori H., Halder S.K., Nonaka Y., Okamoto M.;  
 RT "Cloning of a novel kinase (SIX) of the SNF1/AMK family from high  
 RT salt diet-created rat adrenal."  
 RL FEBS Lett. 453:135-139(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Feldman J.D., Vician L., Crispino M., Hoe W., Baudry M.,  
 RA Herschman H.K.;

RT "The kid2 gene encodes a protein kinase induced by depolarization in brain."  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC SNF1 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AB020480; BAA82673.1; -  
 CC EMBL: AF106937; AAF14191.1; -  
 CC HSSP: P24941; IAQ1.  
 CC InterPro: IPR000719; Euk\_pkinase.  
 CC InterPro: IPR002290; Ser\_thr\_pkinase.  
 CC InterPro: IPR000449; UBA\_domain.  
 CC Pfam: PF00069; pkinase; 1.  
 CC ProDom: PD000001; Euk\_pkinase; 1.  
 CC SMART: SM00165; UBA; 1.  
 CC SMART: SM00220; S\_TKC; 1.  
 CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 CC PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 CC Transferrase; Serine/threonine-protein kinase; ATP-binding.  
 KW DOMAIN 27 278 PROTEIN KINASE.  
 FT NP\_BIND 33 41 ATP (BY SIMILARITY).  
 FT BINDING 56 56 ATP (BY SIMILARITY).  
 FT ACT\_SITE 149 149 BY SIMILARITY.  
 FT CONFLICT 473 473 R -> K (IN REF. 2).  
 SO SEQUENCE 776 AA; 84908 MW; 7BF745AE28F1E6E CRC64;  
 Query Match 19.18; Score 784; DB 1; Length 776;  
 Best Local Similarity 30.58; Pred. No. 1.1e-24;  
 Matches 240; Conservative 104; Mismatches 257; Indels 186; Gaps 28;

DB 463 -----QEPLEGSTGRHRLAEVTHFSPLNPCCIIYSSSAAYSPSEGTSSDCLP 512  
 QY 539 -----SPGGVGGAAMRSLNS-----IRNSFGSPREHRRKQVPAEMSSITPSSP 588  
 DB 513 FSASEGPAGLGGGLAPGLGTSPPVRLASPFGS-----QATPYLQAGLGLATVP 566  
 QY 589 -----ELAKRSWNGNFTSLDKEQIFLVLRKDKPLSIKADI 624  
 DB 567 PVSFQEGRRASDTSITGLIKAFROQLRKNAKRTGFIQLNK----- 606  
 QY 625 VHAFTSPSLSHSVLSQTSFRAEYKASGSPYQKPYRQVDLSSESGEPSPRRGSGG 684  
 DB 607 -----INGLANQV-QQSSIRG-----SRGMSITFHHPA-----PSSGLIG 640  
 QY 665 GGIYSVTFTLLSGPSRRKRYVETI--QAQLLSTHDQPSVQALADEKNGAQTTPAGAPR 742  
 DB 641 -----CTASREGRSLLEVLHQORLLQLQHSNVS--SDVQAQAPLSPV---PY 685  
 QY 743 SLOPPG 749  
 DB 686 VLRPDG 692

# RESULT 7 SNL\_MOUSE ID SNL\_MOUSE STANDARD; PRT; 779 AA.

AC 060670;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable serine/threonine protein kinase SNF1LK (EC 2.7.1.-) (HRT-20)  
 DE (Myocardial SNF1-like kinase).  
 GN SNF1LK OR MSK.  
 OS Mus musculus. (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND REVISIONS TO 8; 16-18; 44-45; 316 AND 435.  
 RC TISSUE=Embryo;  
 RA Ruiz J.C.;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RP [2]  
 RP SEQUENCE OF 1-435 FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=Embryo;  
 RX MEDLINE=95200798; PubMed=7893599;  
 RA Ruiz J.C.; Conlon F.L.; Robertson E.J.;  
 RT Identification of novel protein kinases expressed in the myocardium  
 RT of the developing mouse heart.\*;  
 CC Mech. Dev. 48:153-164(1994).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN LUNG, SKIN, OVARY, HEART AND  
 CC STOMACH. NO EXPRESSION IN BRAIN, LIVER OR SKELETAL MUSCLE.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC SNF1 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U11944; AAA67926.2; -  
 CC HSSP: P24941; IAQ1.  
 CC MGI: MGI:104754; Snf1lk.  
 CC InterPro: IPR000719; Euk\_pkinase.  
 CC InterPro: IPR002290; Ser\_thr\_pkinase.  
 CC InterPro: IPR001245; Tyr\_pkinase.  
 CC InterPro: IPR000449; UBA\_domain.  
 CC Pfam: PF00069; pkinase; 1.  
 CC PRINTS: PR00109; TYRKINASE.  
 CC ProDom: PD000001; Euk\_pkinase; 1.



```

Db 136 FTRQITIGSYCALGIVHDLKPEMLLDHKYIKIADFGMALETEGKLETSQSGSPH 195
QY 196 YACEVIRKEKGYGRADMMSCGVLIFALLVGLALPD--DDNLRQLEKRYGPHNP-- 251
Db 196 YAAPELVISIPVQGFSDVSCVILLFALLTGRLPDEEDGNITLLKLVQGEFEPD 255
QY 252 HTIPPDQGLKMGIEVEPERKLSLEQIOKHMPYLGKHEPDPCLPEAPGR--VAMRSL 309
Db 256 DEISRBAODLIRKILVDEPERIKITRDILKH-----PLLOKVPISRDQSKIRGL 304
QY 310 P-----SNGELDDVLESMASTGCFRDRERLRHRLSEENQEKIYLLIDRK 358
Db 305 PREDIYLPPLSNTSIDTITQNLVILMHGRDPEGIKERLEGANAETLVALLYRFK 364
QY 359 ERYPSCE-----DODLPPRNDVD--PPKRVDSPLMRHGRKRRPERKSMVLSITD 407
Db 365 -----CDYQKELIKQOVKQKQSSISVSVPKRV-----STTPQRRANESL--ISVTS 412
QY 408 AGGGSPPTRALALEMAQHSQSRVSGASTGLSSPLSPSPSP-----VPSFSEPG 460
Db 413 S-----RKKPISFNKTASSASSNLTTPGSSKRLSKNFSSKKKLTIVQSSPTPA 464
QY 461 A 461
Db 465 S 465

```

## RESULT 9

```

SNFL_SCHPO STANDARD; PRT: 576 AA.
AC 074536;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUN-1999 (Rel. 38, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
GN SNFL-like protein kinase (EC 2.7.1.1.).
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Glynnoprez B.,
RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Eger P., Zimmermann W., Medler H., Lehnardt R., Pohl T.M.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Gloux S., Lelaire V., Mottier S.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revelante J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RA "The genome sequence of Schizosaccharomyces pombe.",
RA Nature 415:871-880(2002).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

```

```

CC SNFL SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL031543; CAA20833.1; -
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR002290; Set_thr_Pkinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Euk_Pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_ST; 1.
DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 34 285 PROTEIN KINASE.
FT NP_BIND 40 48 ATP (BY SIMILARITY).
FT BINDING 63 63 ATP (BY SIMILARITY).
FT ACT_SITE 156 156 BY SIMILARITY.
SQ SEQUENCE 576 AA; 65996 MW; E5857E8F171E7B50 CRC64;

```

```

Query Match 18.4%; Score 757.5; DB 1; Length 576;
Best Local Similarity 36.5%; Pred. No. 9,1e-24;
Matches 167; Conservative 95; Mismatches 133; Indels 63; Gaps 12;

```

```

QY 24 PQO--HNAVYPRLEKTIKGGQGLVAVCHTQGVAVAIKIVREKLSSEVL-MKVER 80
Db 22 PPEAIRNHRHIGYPIRREITLGGSGFKVLAHYKQVAKLFTSRQLKSDMHMVER 81
QY 81 EIALIKLIEHPVLAIDVYENKKYLVLEHVSQGEFLDYLRKGRITLPEAKRFFQOI 140
Db 82 EISYLKILRPHIKLVITTPPIDVAVIEY-AGGELFDYIVEKKNTEDEGRFFQOI 140
QY 141 VSALDPCHSISICRDLKPEMLLDKKNIRIADFGMASLOVGDLSLFTSCGSPHYAPE 200
Db 141 ICALIEYCHRRKIYHRDKPEMLLDMLNVAIDFGLSNIMTDGFLTSCGSPHYAPE 200
QY 201 VIKGEKGRADMMSCGVLIFALLVGLALPDNDNLQLEKRYGPHNPHTPPDQCS 260
Db 201 VINGKILYAGEVDVSCGVLIVLVRGLRPPDDEFIPLEKRVNSCYVMPDPLSPGAS 260
QY 261 LIRGMIEVEPERKLSLEQIOKHMPYLGKHEPDPCLPEAPGRVAMRSLPSNGEILPD 318
Db 261 LIRRMIVADPMQRIITQEIIRDPWF--NVNLPDLRMEVQGSYADSRVSKGE---- 314
QY 319 VLESMASTGCFRDRERLRHRLSEENQEKIYLLIDRKRRYPSCEDQLPPRNDVDP 378
Db 315 -----AMGFSED--YIVEALRSDENNEVEKAYNL-----HENOVLOEKSHLS-K 356
QY 379 RKRVD-----PMLSRHGRRRPERKSMVLSITDAGCGS----- 413
Db 357 SKRVDLSVSPAPSEYSEIQKSKQKQELIDPLBEPNRTVSPPIYAQTIDSNICVL 416
QY 414 -VYPTRALALEMAQHSQSRVSGASTGLSSPLSPSPS 450
Db 417 VPTAEKKNKLEM-----RTLADAAAGAVDSQSTRKKS 447

```

## RESULT 10

```

SNFL_YEAST STANDARD; PRT: 633 AA.
AC P06782;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Carbon catabolite derepressing protein kinase (EC 2.7.1.1.).

```

GN SNF1 OR CARI OR CCRI OR PAS14 OR GIC2 OR YDR477M OR D8035.20.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86289463; PubMed=3526554;  
 RA Celenza J.L., Carlson M.;  
 RT "A yeast gene that is essential for release from glucose repression  
 RT encodes a protein kinase."  
 RL Science 233:1175-1180(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,  
 RA Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,  
 RA Huntke-Smith S., Hymen R., Komp C., Laskari D., Lew H., Lin D.,  
 RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,  
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,  
 RA Winant A., Yelton M., Botstein D., Davis R.W.;  
 RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 274-284; 528-539 AND 622-630, AND PHOSPHORYLATION SITE.  
 RX MEDLINE=94131988; PubMed=7905477;  
 RA Mitchellhill K.I., Stapleton D., Gao G., House C., Mitchell B.,  
 RA Katsis F., Witters L.A., Kemp B.E.;  
 RT "Mammalian AMP-activated protein kinase shares structural and  
 RT functional homology with the catalytic domain of yeast Snf1 protein  
 RT kinase."  
 RL J. Biol. Chem. 269:2361-2364(1994).  
 CC -1- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT  
 CC INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY  
 CC PROTEIN SNF1. INTERACTS ALSO WITH SIPL, SIPL2 AND GAL83. COULD  
 CC PHOSPHORYLATES CAT8.  
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC SNF1 SUBFAMILY.  
 CC -----  
 CC This SNF1-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; M13871; AAA35058.1; -;  
 CC DR EMBL; U33050; AAB64904.1; -;  
 CC DR PIR; A26030; A26030.  
 CC DR HSSP; P24941; 1HCL.  
 CC DR SGD; S0002885; SNF1.  
 CC DR InterPro; IPR000719; Euk\_pkinase.  
 CC DR InterPro; IPR002290; Ser\_thr\_Pkinase.  
 CC DR Pfam; PF00069; Pkinase; 1.  
 CC DR ProDom; PD000001; Euk\_Pkinase; 1.  
 CC DR SMART; SM00220; S\_TKC; 1.  
 CC DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 CC DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC DR Transferrase; Serine/threonine-protein kinase: ATP-binding;  
 CC Phosphorylation; Carbohydrate metabolism; Nuclear protein.  
 CC KM  
 CC FT DOMAIN 18 32 POLY-HIS.  
 CC FT DOMAIN 55 306 PROTEIN KINASE.  
 CC FT NP\_BIND 61 69 ATP (BY SIMILARITY).  
 CC FT BINDING 84 84 ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 177 177 BY SIMILARITY.  
 CC FT MOD\_RES 210 210 PHOSPHORYLATION (AUTO-).  
 CC FT MOD\_RES 210 210 F5C6335C5C986C4E3 CR664;  
 CC FT SEQUENCE 633 AA; 72045 MW; F5C6335C5C986C4E3 CR664;  
 CC  
 CC Query Match 18.3%; Score 753.5; DB 1; Length 633;  
 CC Best local similarity 32.8%; Pred. No. 1.4e-23;  
 CC Matches 166; Conservative 103; Mismatches 144; Indels 93; Gaps 14;

14 AYHLRPHRPPPOHA-----YGPRLKTLTGKGTGLKLVNCT 55  
 Db SHHHHHHHHHHHHGGSGNSTLNPKSLADGHHIGNYQIVKTLGSGSGKVLAVHTT 76  
 QY 56 TGQKVALIKYNNREKLSVLM-KVERELATIKLEHGHVLYKLHGVYENKVKLYLVLEHS 114  
 Db 77 TGQKVALIKYNNREKLSVLM-KVERELATIKLEHGHVLYKLHGVYENKVKLYLVLEHS 114  
 QY 115 GGELFDLYKKGRITPKRKARFQRIVSALDFCHSISCHRDLPENILDEKNRIAD 174  
 Db 136 GNELPDIYVQDKASDEQARFPQIIISAVYCHRHKLIVHNDLPENILDEHNLVLTAD 195  
 QY 175 FGMSLQVDSLETSCTGSPHYACPEVYKGEKYGRADAMSCGYIFALLVGLPDPD 234  
 Db 196 FGLSNIMTDGNEFLKTSCTGSPHYACPEVYKGEKYGRADAMSCGYIFALLVGLPDPD 255  
 QY 235 NLRLQLEKVKGVHMPFIPDPDQSLRGMEVEPEKLSIQOKRPVLYGKHE--- 291  
 Db 256 SIPVLEKISNGVYTLPEFLSPGAAGLIKRLIYPLNRIHSEIMODDFKVDLPETLL 315  
 QY 292 -PDCLEPAPGRVAMR-----SLPSNGELDPVLESMAISGCFRDRRLHRLSEEE 344  
 Db 316 PPD-LKHPPEENNNNSKDGSSPDNDIEDNLVNLISPMGY-EKDEIYESLESSED 372  
 QY 345 ---NOEMITYLLIDRKERTPSCDQDLPFRNDY-----DPP-----RK 380  
 Db 373 TPAPNEIDAYMLIKENKSLI-----KDMKANKSVSDEIDTFLSPPFPQOQSKSHQS 427  
 QY 381 RFDSPMLSRHGRKRPERSKMEVLSITDAGGSGSPPTPRALE---MQHOSRSKSVGA 436  
 Db 428 QVDHETAKOHARM-----ASATIQGRTHYQSPFMQYKREDSVSLT 470  
 QY 437 STGL-----SSPLSPRSV 452  
 Db 471 PTLSPQIRANMLAQSPASKISPL 496  
 CC  
 CC RESULT 11  
 CC SNF1\_CANTR STANDARD; PRT; 619 AA.  
 CC ID 094168;  
 CC AC 15-JUL-1999 (Rel. 38, Created)  
 CC DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 CC DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 CC DE Carbon catabolite derepressing protein kinase (EC 2.7.1.-).  
 CC GN SNF1.  
 CC OS Candida tropicalis (Yeast).  
 CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC CC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 CC OX NCBI\_TaxID=5482;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RA Kanai T., Ogawa K., Ueda M., Tanaka A.;  
 CC RT "Genetic evaluation of the function of SNF1 in Candida tropicalis."  
 CC RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 CC CC -1- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT  
 CC INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY  
 CC PROTEIN SNF4. COULD PHOSPHORYLATES CAT8 (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC SNF1 SUBFAMILY.  
 CC -----  
 CC This SNF1-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; AB024535; BAA75889.1; -;  
 CC DR HSSP; Q63450; 1A06.

DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR Pfam: PF00069; pkinase.1.  
 DR ProDom: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00220; S\_TKC.1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP.1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST.1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM.1.  
 DR Transferrase: Serine/threonine-protein kinase; ATP-binding;  
 KW Phosphorylation; Carbohydrate metabolism; Nuclear protein.  
 FT DOMAIN 17 29 POLY-HIS.  
 FT NP\_BIND 52 303 PROTEIN KINASE.  
 FT BINDING 58 66 ATP (BY SIMILARITY).  
 FT ACT\_SITE 81 81 ATP (BY SIMILARITY).  
 FT MOD\_RES 174 174 BY SIMILARITY.  
 FT MOD\_RES 207 207 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 SQ SEQUENCE 619 AA; 70323 MW; 0FC1FC3DCE706D7 CRC64;

Query Match 17.8%; Score 730; DB 1; Length 619;  
 Best Local Similarity 37.8%; Pred. No. 1,2e-22;  
 Matches 153; Conservative 90; Mismatches 118; Indels 44; Gaps 10;

QY 15 YLPHPHPHPOHAQ-----YVPRLEKTLGKGTGLVYHCITGQVAVI 62  
 DB 21 HHHHHHHHSDPAPCPIDPNVNPANRIGRGQIKTLGSGSGKVLADHVGIGQVAL 80  
 QY 63 KLVNREKSESLVM-KVEREIALKLEPHVYLKLDYENKKYLYLHVSGGELEFDY 121  
 DB 81 KLVNREKSESLVM-KVEREIALKLEPHVYLKLDYENKKYLYLHVSGGELEFDY 121  
 QY 122 LVKRGRLPPEAKRFERQVLSALDFCHSYSGHRLKPEVLLDEKNNITADPGMSIQ 181  
 DB 140 IVORGEKMPEDARFEPQOITAAVEYGHRRKIYARDLKPEVLLDDQVNAVADPGMSINM 199  
 QY 182 VGDLSLESCGSPHACPRVYIGKGTGRADMSGCVLLFALLGALPPDDNLROLLE 241  
 DB 200 TDGNLTKTSCGSPNAAPEVIGSKYAGPEVYVSSGVLLYVLCRLPPDDEFIPALFK 259  
 QY 242 KVRKGVFMPHPIPPDCOSILRGMIETVEPEKRLSLQIOKHPIYLGKHEHDDCIEPAPG 301  
 DB 260 KISNGVYLPNPLSGAKHLLRLMVLVNPANRIGRGQIKTLGSGSGKVLADHVGIGQVAL 315  
 QY 302 RRVAKRSLPSNGEILDVLESMA-SLGCERDR-----ERLHRE-----LISEENOE 347  
 DB 316 --LSKITSKIDIDEDVIALSVYMGYDRDEITISYERKANRBAAGAPPTQSKSTNE 372  
 QY 348 KMYVLLDRKERYPSCHDOLPPRNDV-----PPKRVDSPP 385  
 DB 373 VLDAYLLM--KENHTLVYDLKKSKSENIESTLSPPSSSFPNP 415

RESULT 12  
 AAKI\_HUMAN STANDARD; PRT: 550 AA.  
 AC G13131; 000286; GQUNQ4;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 5'-AMP-activated protein kinase, catalytic alpha-1 chain (EC 2.7.1.-)  
 GN (AMPK alpha-1 chain).  
 DE PRKAA1 OR AMPK1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Mammary gland;  
 RA Yano K.;  
 RT "Nucleotide sequence of cDNA for human AMP-activated protein kinase  
 alpha-1.";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC TISSUE-Blood;  
 RX MEDLINE=20499367; PubMed=11042152;  
 RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,  
 RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,  
 RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;  
 RT "Cloning and functional analysis of cDNAs with open reading frames for  
 RT 300 previously undefined genes expressed in CD34+ hematopoietic  
 RT stem/progenitor cells.";  
 RL Genome Res. 10:1546-1560(2000).  
 RN [3]  
 RP SEQUENCE OF 27-200 FROM N.A.  
 RC TISSUE-Intestine;  
 RA Tabada E.N., Hickey D.A.;  
 RL Submitted (Apr-1995) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 294-550 FROM N.A.  
 RC TISSUE-Liver;  
 RX MEDLINE=96132781; PubMed=8557660;  
 RA Stapleton D., Mitchell H.I., Gao G., Widmer J., Mitchell B.J.,  
 RA Teh T., House C.M., Fernandez C.S., Cox T., Witters L.A.,  
 RA Kemp B.E.;  
 RT "Mammalian AMP-activated protein kinase subfamily.";  
 RL J. Biol. Chem. 271:611-614(1996).  
 CC -1- FUNCTION: RESPONSIBLE FOR THE REGULATION OF FATTY ACID SYNTHESIS  
 CC BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO REGULATES  
 CC CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND INACTIVATION OF  
 CC HORMONE-SENSITIVE LIPASE AND HYDROXYMETHYLGLUTARYL-COA REDUCTASE.  
 CC APPEARS TO ACT AS A METABOLIC STRESS-SENSING PROTEIN KINASE  
 CC SWITCHING OFF BIOSYNTHETIC PATHWAYS WHEN CELLULAR ATP LEVELS ARE  
 CC DEPLETED AND WHEN 5'-AMP RISES IN RESPONSE TO FUEL LIMITATION  
 CC AND/OR HEPFOXIA. THIS IS A CATALYTIC SUBUNIT (BY SIMILARITY).  
 CC -1- SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT (BY SIMILARITY).  
 CC NON-CATALYTIC SUBUNITS  
 CC -1- PIM: AUTOPHOSPHORYLATED (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC SNF1 SUBFAMILY.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL: AB022017; BAA36547.1; -;  
 DR EMBL: AF100763; AAD43027.1; -;  
 DR EMBL: U22456; AAA64850.1; -;  
 DR EMBL: Y12856; CAAT3361.1; -;  
 DR HSSP: Q63450; IA06.  
 DR GeneW: HGNC:9376; PRKAA1.  
 DR MIM: 602739; -;  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR Pfam: PF00069; pkinase.1.  
 DR ProDom: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00220; S\_TKC.1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP.1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST.1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM.1.  
 KW Transferrase: Serine/threonine-protein kinase; Fatty acid biosynthesis;  
 KW Phosphorylation; ATP-binding.  
 FT DOMAIN 18 270 PROTEIN KINASE.  
 FT NP\_BIND 32 32 ATP (BY SIMILARITY).  
 FT BINDING 47 47 ATP (BY SIMILARITY).  
 FT ACT\_SITE 141 141 BY SIMILARITY.  
 FT MOD\_RES 174 174 PHOSPHORYLATION (AUTO-) (POTENTIAL).  
 FT CONFLICT 28 28 T -> A (IN REF. 3).  
 FT CONFLICT 133 133 A -> V (IN REF. 3).  
 FT CONFLICT 199 199 I -> L (IN REF. 3).  
 FT CONFLICT 260 260 S -> T (IN REF. 2).  
 SQ SEQUENCE 550 AA; 62793 MW; 3316183D744DE325 CRC64;

Query Match 17.6%; Score 725; DB 1; Length 550;  
 Best Local Similarity 40.3%; Pred. No. 1.7e-22;  
 Matches 154; Conservative 67; Mismatches 115; Indels 46; Gaps 7;

```

OY 31 VGPRLKLTGKGTGLVGLVHCITGOKVAIKIVNREKL-SESYLMKVEREITALKLE 89
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 15 IGHVILDTLVGTFGFKVKGKHELTGKVAIKILNRQKIRSLDVGKIRREIQMLKLR 74
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 90 HPAVLKLDVYENKRYLYLVLEHVSGETLPDYLVKGRITREAKRFQOYSAIDFGCS 149
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 75 HPHIKLYQVYISPSDIFWMEVYSGELFDYICKNGRLDERESRLFOQLISGVYDCHR 134
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 150 YSICHRDLKPEKNLLDEKNNIRIADFGMASLOVGDLSLETSCGSPHYACPEYIKGEKYDG 209
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 135 HMYVHRDLKPEKNLLDAMNAKIDPGLSNMMSDGEFLRTSCGSPHYAARVYISGRLYAG 194
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 210 RRADMVSGVILFALLVGLALPDDNLRQLLEKVRGVFHPHFIPPCQSLRGMIYE 269
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 195 PEVDIMSSGVILYALCGLTFPDDHVPFLFKICDGIFFYPOYINPVSISLKHMLQVD 254
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 270 PEKRLSLEQIOKHFWYLGKHEPDCLEPARGRYAMRSLPNGELDPVLESMAISGCF 329
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 255 PMKRASIKDIREHEWF---KODLPKYLEPE-----DYSYSTMT----- 290
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 330 RDRERLRELRSEENOEKMI-----YYLLDRKRRYSCEDQDLPFRND 374
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 291 -DDEALKEVCEKEFSCSEEVYSLCYLNRRHODPLAVAYHLLIDNRRLMNEAKDFYLA--- 345
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 375 VDPKRVDSPLSR-HGKRRP 395
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 346 TSPDPSFLDHLHTRPHERVP 367
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  
```

## RESULT 13

```

AKL1_RAT STANDARD; PRT; 548 AA.
ID AKL1_RAT
AC P54645;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 5'-AMP-activated protein kinase, catalytic alpha-1 chain (EC 2.7.1.-)
DE (AMPA alpha-1 chain).
GN PRKAA1 OR AMPK1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid-10116;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN-Sprague-Dawley; TISSUE-Liver, and Hypothalamus;
RX MEDLINE=96133781; PubMed=8557660;
RA Stapleton D., Mitchell K.I., Gao G., Widmer J., Mitchell B.J.,
RA Teh T., House C.M., Fernandez C.S., Cox T., Witters L.A.,
RA Kemp B.E.;
RA "Mammalian AMP-activated protein kinase subfamily.";
RT "Mammalian AMP-activated protein kinase subfamily.";
RL J. Biol. Chem. 269:29343-29346(1994).
-1- FUNCTION: RESPONSIBLE FOR THE REGULATION OF FATTY ACID SYNTHESIS
BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO REGULATES
CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND INACTIVATION OF
HORMONE-SENSITIVE LIPASE AND HYDROXYMETHYLGLUTARYL-COA REDUCTASE.
APPEARS TO ACT AS A METABOLIC STRESS-SENSING PROTEIN KINASE
SWITCHING OFF BIOSYNTHETIC PATHWAYS WHEN CELLULAR ATP LEVELS ARE
DEPLETED AND WHEN 5'-AMP RISES IN RESPONSE TO FUEL LIMITATION
  
```

AND/OR HYPOXIA. THIS IS A CATALYTIC SUBUNIT.  
 -1- SUBUNIT: HETEROTRIMER OF A CATALYTIC SUBUNIT, A BETA AND A GAMMA  
 NON-CATALYTIC SUBUNIT.  
 -1- TISSUE SPECIFICITY: LOW EXPRESSION IN KIDNEY, LIVER, LUNG, HEART,  
 AND BRAIN.  
 -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 SNF1 SUBFAMILY.

-----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; UA0819; AAC52355.1; -  
 CC HSSP; Q63450; 1A06  
 CC InterPro; IPR000719; Euk\_pkinase.  
 CC InterPro; IPR002290; Ser\_thr\_pkinase.  
 CC Pfam; PF00069; pkinase; 1.  
 CC ProDom; PD000001; Euk\_pkinase; 1.  
 CC SMART; SM00220; S\_TKc; 1.  
 CC DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 CC DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC KW Transferase; Serine/threonine-protein kinase; Fatty acid biosynthesis;  
 CC KW Phosphorylation; ATP-binding; Multigene family.  
 CC FT DOMAIN 16 268 PROTEIN KINASE.  
 CC FT NP\_BIND 22 30 ATP (BY SIMILARITY).  
 CC FT BINDING 45 45 ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 139 139 BY SIMILARITY..  
 CC FT MOD\_RES 172 172 PHOSPHORYLATION (AUTO-) (POTENTIAL).  
 CC FT SEQUENCE 548 AA; 62599 MW; 50CA3281C195F867 CRC64;

Query Match 17.6%; Score 722; DB 1; Length 548;  
 Best Local Similarity 40.1%; Pred. No. 2.2e-22;  
 Matches 153; Conservative 68; Mismatches 115; Indels 46; Gaps 7;

```

OY 31 VGPRLKLTGKGTGLVGLVHCITGOKVAIKIVNREKL-SESYLMKVEREITALKLE 89
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 13 IGHVILDTLVGTFGFKVKGKHELTGKVAIKILNRQKIRSLDVGKIRREIQMLKLR 72
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 90 HPAVLKLDVYENKRYLYLVLEHVSGETLPDYLVKGRITREAKRFQOYSAIDFGCS 149
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 75 HPHIKLYQVYISPSDIFWMEVYSGELFDYICKNGRLDERESRLFOQLISGVYDCHR 132
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 150 YSICHRDLKPEKNLLDEKNNIRIADFGMASLOVGDLSLETSCGSPHYACPEYIKGEKYDG 209
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 133 HMYVHRDLKPEKNLLDAMNAKIDPGLSNMMSDGEFLRTSCGSPHYAARVYISGRLYAG 192
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 210 RRADMVSGVILFALLVGLALPDDNLRQLLEKVRGVFHPHFIPPCQSLRGMIYE 269
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 193 PEVDIMSSGVILYALCGLTFPDDHVPFLFKICDGIFFYPOYINPVSISLKHMLQVD 252
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 270 PEKRLSLEQIOKHFWYLGKHEPDCLEPARGRYAMRSLPNGELDPVLESMAISGCF 329
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 253 PMKRATIDIREHEWF---KODLPKYLEPE-----DYSYSTMT----- 288
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 330 RDRERLRELRSEENOEKMI-----YYLLDRKRRYSCEDQDLPFRND 374
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 289 -DDEALKEVCEKEFSCSEEVYSLCYLNRRHODPLAVAYHLLIDNRRLMNEAKDFYLA--- 343
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 375 VDPKRVDSPLSR-HGKRRP 395
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 344 TSPDPSFLDHLHTRPHERVP 365
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  
```

## RESULT 14

```

KCC4_YEAST STANDARD; PRT; 915 AA.
ID KCC4_YEAST
AC P25389; P87005;
DT 01-MAY-1992 (Rel. 22, Created)
  
```

DT 30-MAY-2000 (Rel. 39, last sequence update)  
 DE 30-MAY-2000 (Rel. 39, last annotation update)  
 DT Probable serine/threonine-protein kinase YCL24W (EC 2.7.1.1-).  
 GN YCL24W OR YCL24W  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
 OX NCBI\_TaxId=4932;  
 RN [1]  
 RP SEQUENCE OF 1-569 FROM N.A.  
 RA Duesterhoft A., Erdmann D., Hegemann J., Philippson P.,  
 RA Schweitzer B., Spiegelberg R.;  
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 567-915 FROM N.A.  
 RA Oliver S.G., Anwar R., Brown A., Gent M.E., Indge K.J., James C.M.,  
 RA Staveva L.I.;  
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP REVISIONS.  
 RA Gromada R.;  
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC NIM1 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X59720; CAA42361.1;  
 CC DR PIR; S19351; S19351.  
 CC DR HSSP; Q63450; 1A06.  
 CC DR SGD; S0000529; YCL024W.  
 CC DR InterPro: IPR000719; Euk\_pkinase.  
 CC DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 CC DR InterPro: IPR001245; Tyr\_pkinase.  
 CC DR Pfam; PF00069; pkinase; 1.  
 CC DR PRINTS; PR00109; TYRKINASE.  
 CC DR PRODOM; PD000001; Euk\_pkinase; 1.  
 CC DR SMART; SM00220; S\_TKC; 1.  
 CC DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 CC DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC DR Hypothetical protein; Transferase; Serine/threonine-protein kinase;  
 CC ATP-binding.  
 CC KW DOMAIN 21 285 PROTEIN KINASE.  
 CC FT NP\_BIND 27 35 ATP (BY SIMILARITY).  
 CC FT BINDING 50 50 ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 152 152 BY SIMILARITY.  
 CC FT SEQUENCE 915 AA; 102688 MW; BFB01C8CA43AC181 CRC64;  
 SQ  
 Query Match 17.6%; Score 722; DB 1; Length 915;  
 Best Local Similarity 38.3%; Pred. No. 3.5e-22;  
 Matches 171; Conservative 80; Mismatches 155; Indels 40; Gaps 11;  
 QY 31 VGPRLKLTGKGTGLVKGAVHCITGKQVAKIVNREKLS-----ESVL-MKERE 81  
 DB 18 IGPMLKLTGKGTGKQVLAQHRTGHTAVKISKLFNNNGNSNDSDVLPYNIERE 77  
 QY 82 IALLKLHPVLAHVDYENKKYLYLVLEHVSGLLEFLYKKGRLTPKARKFFQIV 141  
 DB 78 IVIKLLSHPNVLSLYDWTNNNNYLLILEYAEKGLFVLVDHGPLPREEAICFRQII 137  
 QY 142 SALDCHSYICHRDLPENLLDEKNNIRIADFGMASLYOVDSILETSCGSHYACPEV 201  
 DB 138 IGISYCHALGIYHRLKRENLLDSFYNIKIDFGNALQTDADLETSCGSHYAPET 197  
 QY 202 IKGKYDGRADMWSCGVLVALLGALPFDD--NLRLLEKVKGVHMPH--FIPDP 257

DB 198 VSLPYEGFASDVWSCGVILFALLTGLRPPDEENGVRDILLKQKGFEMPNDTEISR 257  
 QY 258 CQSLGMEIYEVEERKSLSEIQKHRYLGCKHE---PDCLEPAPRRVAMRSLP---- 310  
 DB 258 ADELKILVWDERQRIKINDILSH--LLKKYQTIKDSKSIKLPRENTIYLPADSNN 315  
 QY 311 -SNGELDPVLEEMASLGCGRDERRLHRESEBNEKMIYIYLLRKERYPCCEODL 369  
 DB 316 HTSASIDDSITLQNLVYLMHGRHADDIYSKLENGTNEKITLITALLY--REKLSVRSNK 373  
 QY 370 PPRNDVPPRRVDSFMLSNGRRRERKSMELYSLTDAGGSPVPTRALMAHQSR 429  
 DB 374 KNRNKKIKTKKRSSTLSSSSSLNNRSLQ-----STPRRTSK--RRSRE 419  
 QY 430 -SRVSAGASTGLSSSPSPSPVFS 454  
 DB 420 FSSSRKRSFLSSNPTDSSPIPLRS 445  
 RESULT 15  
 ID SNF1\_CANGA STANDARD; PRT; 611 AA.  
 AC 000372;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, last sequence update)  
 DT 15-JUL-1999 (Rel. 38, last annotation update)  
 DE Carbon catabolite derepressing protein kinase (EC 2.7.1.1-).  
 GN SNF1  
 OS Candida glabrata (Yeast) (Torulopsis glabrata).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxId=5478;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCCLS84;  
 RX MEDLINE=97101049; PubMed=8945576;  
 RA Peltier R., Kwon-Chung K.J.;  
 RT "Disruption of the SNF1 gene abolishes trehalose utilization in the  
 RT pathogenic yeast Candida glabrata."  
 RL Infect. Immun. 64:5269-5273(1996)  
 CC -1- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT  
 CC INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY  
 CC PROTEIN SNF4. COULD PHOSPHORYLATES CAT8 (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC SNF1 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L78130; AAB48642.1;  
 CC DR HSSP; P24941; 1HCL.  
 CC DR InterPro: IPR000719; Euk\_pkinase.  
 CC DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 CC DR Pfam; PF00069; pkinase; 1.  
 CC DR PRODOM; PD000001; Euk\_pkinase; 1.  
 CC DR SMART; SM00220; S\_TKC; 1.  
 CC DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 CC DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
 CC Phosphorylation; Carbohydrate metabolism; Nuclear protein.  
 CC KW DOMAIN 6 17  
 CC FT NP\_BIND 39 290 PROTEIN KINASE.  
 CC FT BINDING 45 53 ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 161 161 ATP (BY SIMILARITY).

FT	MOD_RES	194	194	PHOSPHORYLATION (AUTO-) (BY SIMILARITY)
SEQ	SEQUENCE -	611 AA;	70049 MW;	89E17812A4900CD0 CRC64;

FT	MOD_RES	194	194	PHOSPHORYLATION (AUTO-) (BY SIMILARITY)
SEQ	SEQUENCE -	611 AA;	70049 MW;	89E17812A4900CD0 CRC64;

Query Match	17.58; Score 720; DB 1; Length 611;
-------------	-------------------------------------

Matches 148; Conservative 97; Mismatches 129; Indels 42; Gaps 10;

Matches 148; Conservative 97; Mismatches 129; Indels 42; Gaps 10;

```

QY 15 YHLPHPHPPOHNOY-----VGPRLEETLKGOGTGVLGVNHTIQOKYA 61
D 7 HHHHHHHHHHNSNSTYSNKVSYSLADSGSRGNGQYKTLTGESSFQKYLAVHTTQOKYA 66
QY 62 IKIVNRKLESSEYLM-KVEBEIALIKLIEHPHVLKLDVYENKKYLYLVEHVSGELED 120
D 67 LKSINKKVLAKMQGRIIDREISYTLRLRPHILIKLYDAVYSKDEILIMVEY-AGNELED 125
QY 121 YLYAKGRUTPEAKRFROLYSALDFCHSISICHROLKRPENLLDDEKNNIRIDPFQMASL 180
D 126 YLYVRNKKMSDEBARFPOOIIISAEYCHRRKIVYHRDLKRPENLLDDEHLVYK IADFGISLNI 185
QY 181 QVGSLLTSGSGSHYACPEYIKCEKVDGRRADWMSGVTLLFALYCALPFDDDNIRQLL 240
D 186 MTDGNGFLKTSGSGSNYNAPEYISKLTYAGEVDWMSGVTLLYMLCRRLPFDDDESTIPLYE 245
QY 241 EKVARGVFMHMFTRPPDCQSLNLRGMEVEBERKLSLQIQKHWPYLGKKEHDPCELPAP 300
D 246 KNISGVYTLTKPLSPGASDLIKMLLVNPLNRISIHIEIMDEWF---KVDLAEVLYPD 302
QY 301 GRVYAMNSLPBNG-----ELDDPYLESMA-SLGCFRERKRLHRLHRSLEE-----NOEKMI 350
D 303 LKOOEOPFKKSGNEENVEIDEXVYTLISKTM--YDKDEIYALLESSEDTPAYNEIRNA 360
QY 351 YLLLLDRKERAPCED-----QDL-----PRANDVDPKRVDSFMLRSHGRKR 394
D 361 YLLINDNSLIKMDKONNTVQELDTLTSOSPFPFOONGMKAKESEQKKKHSRR 416

```

Db 7 HHHHHHHHNSGSYVSNKVSSLDGSRVGN YQIVKTLGEGSF GKVKLAIHVT TGQKVA 666

62 IKIVNREKLSVLM-KVEREIAIKLIEHPVLKLDVYENKKYLLVLEHVS GGELED 120

Db 67 LKSINKVLAQSDMGRIDREISYLRLLRPHI IKLYDVIKSKDEIIMVIEY-AGNELEFD 125

121 YLVKKGR LTPKEARKFFRQIVSALDFCHSY SICH RDLKPENLLDEKNNIRIADFGMASL 180

Db 126 YIVQRNKMSEQEARRFQOIISAVEYCHRHKI VHRDLKPENLLLDEHLNVKIADEGLSNI 185

```
QY      181 QVGD S L E T S C G S P H Y A C P E V I K G E K Y D G R R A D M W S C V I L F A L L V G A L P F D D D N L R Q L L    240
```

Db 186 MTDGNFLKTS CGSPNYA AFEV ISGKLYAGPEVDVWSCGVILYVMLCRRLPFDDESIPLF 245

241 EKVKRGVFHMPHFIPPPDCQSLRGMIEVEPERKRLSLEQIQKHPWYLGCKHEPDPCLLEPAP 300

Db 246 KNISNGVYTLPKFLSPGASDLIKRMLIVNPLNRISIHEIMQDEWF---KVDLAEYLVpQD 302

```
QY      301 GRVAMRSLPSNG-----ELDPVLSEMA-SLGCFRDRERLHRELKSEE---NQEKMI    350
```

Db 303 LKQQEQFNKSGNEENEVEEIDDEMVTLSKTMG--YDKDEIYEALSESSEDTPAYNEIRNA 360

```

351 YLLDRKERYPSCED-----QDL-----PRRNDVDPKRKRVDSPMLSRHGKRR 394
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY

```

Db 361 YILIKDNKSLIKDMKQDNNVTQELDTFLSQSPPTFEQONGDGMKASEDQKKKHSGR 416

Search completed: April 16, 2003, 12:30:21  
Job time : 34 secs

Job time : 34 secs



XX

CC Nucleic acids encoding human kinase polypeptides, useful for preventing  
 CC diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
 CC neuronal-associated diseases, and microbial infections -  
 CC Claim 7, Figure 2; 433pp; English.  
 CC AA003501-AA003557 represent novel human protein kinases #1-57. The  
 CC novel protein kinases have been identified as members of the tyrosine  
 CC serine/threonine kinase (PTK and STK) families. The polynucleotides  
 CC encoding protein kinases and the polypeptides may be used in the  
 CC prevention, diagnosis and treatment of diseases associated with  
 CC inappropriate kinase expression. For example, they may be used to treat  
 CC cancers (especially cancers of haematopoietic origin), cardiovascular  
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
 CC immune related diseases (e.g. rheumatoid arthritis), neurological  
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
 CC Additionally, polynucleotides encoding protein kinases may be  
 CC used for gene therapy and as DNA probes in diagnostic assays.  
 CC The protein kinase polypeptides may be used as antigens in the production  
 CC of antibodies against the protein kinases and in assays to identify  
 CC modulators of protein kinase expression and activity.

CC Sequence 794 AA:

Query Match 96.38; Score 3956.5; DB 22; Length 794;  
 Best Local Similarity 98.14; Pred. No. 2.2e-255;  
 Matches 756; Conservative 1; Mismatches 9; Indels 5; Gaps 1;

QY 8 GGGGSPAYHLPHRPHROHAYGVYRLETKGQNGVILGVHCTGOKVAIKIVNR 67  
 DB 29 GGGG-----EEEAEEGRHAAQVYGVYRLETKGQNGVILGVHCTGOKVAIKIVNR 83  
 QY 68 EKLSSEVLMKVEREIALTKLEHPVVLKLDHYENKKYLYLVLEHVSSEGEDEYLVKKGR 127  
 DB 84 EKLSSEVLMKVEREIALTKLEHPVVLKLDHYENKKYLYLVLEHVSSEGEDEYLVKKGR 143  
 QY 128 LTPKEARKFPRQIYASALDFCHSYISICHDLKPENLIDKNNIRIADFGMSLQVDSLL 187  
 DB 144 LTPKEARKFPRQIYASALDFCHSYISICHDLKPENLIDKNNIRIADFGMSLQVDSLL 203  
 QY 188 ETSGSGPHYACPEYIKGEKYDKGRADMMSCGVILFALLVGAIPFDNDNRKOLEKVRG 247  
 DB 204 ETSGSGPHYACPEYIKGEKYDKGRADMMSCGVILFALLVGAIPFDNDNRKOLEKVRG 263  
 QY 248 FHMHPHPIPPDCQSLIKGMIEVEPEKRLSLQIOLKHPVYLGGKHEPDCLEPAGRRVNR 307  
 DB 264 FHMHPHPIPPDCQSLIKGMIEVEPEKRLSLQIOLKHPVYLGGKHEPDCLEPAGRRVNR 323  
 QY 308 SLPSNGELDDVYLESMASLGCGRDRERHRELRSSENOEKMTYLLIDRKERYSCEDQ 367  
 DB 324 SLPSNGELDDVYLESMASLGCGRDRERHRELRSSENOEKMTYLLIDRKERYSCEDQ 383  
 QY 368 DLPPRDNDVDPKRRKVDSPMLSRHGRKRPERSMEVLSITDAGGGSPVTRALAEAOHS 427  
 DB 384 DLPPRDNDVDPKRRKVDSPMLSRHGRKRPERSMEVLSITDAGGGSPVTRALAEAOHS 443  
 QY 428 QRSRSVSGASTGLSSSPLSPSPVFSFSPGAGDEARGGSPSTKQTLPSPRGRRGG 487  
 DB 444 QRSRSVSGASTGLSSSPLSPSPVFSFSPGAGDEARGGSPSTKQTLPSPRGRRGG 503  
 QY 488 AGQPPPPPSKRSRPLPPPPSGSSGTPPLHSLPHTPASPGTGTTPPPSPGGGVGA 547  
 DB 504 AGQPPPPPSKRSRPLPPPPSGSSGTPPLHSLPHTPASPGTGTTPPPSPGGGVGA 563  
 QY 548 AMRSRLNSTINSFSGSPRFRHKKQVPAEEMSSLTPESSPELARSMWGNFISLDKEQ 607  
 DB 564 AMRSRLNSTINSFSGSPRFRHKKQVPAEEMSSLTPESSPELARSMWGNFISLDKEQ 623  
 QY 608 IFVLVADKPLSLIKADIVHAFSLSPSLSHSVLSQTSFRAEKASGSPVFOQPVRRQVDI 667  
 DB 624 IFVLVADKPLSLIKADIVHAFSLSPSLSHSVLSQTSFRAEKASGSPVFOQPVRRQVDI 683

QY 668 SSSEGPSPRRDGGGGGCIYVTFLLISGSPRRKRVETIOAQLISTHDQPSVALAD 727  
 DB 684 SSSEGPSPRRDGGGGGCIYVTFLLISGSPRRKRVETIOAQLISTHDQPSVALAD 743  
 QY 728 EKNGAOTRPAAGAPPRLPPPGRRDPELSSPPRGPKDKKLLATNTCTPLP 778  
 DB 744 EKNGAOTRPAAGAPPRLPPPGRRDPELSSPPRGPKDKKLLATNTCTPLP 794  
 RESULT 2  
 ID AAE16271 standard; Protein; 794 AA.  
 XX AAE16271;  
 AC AAE16271;  
 DT 26-MAR-2002 (first entry)  
 XX Human kinase PKIN-17 protein.  
 DE Human kinase PKIN-17 protein.  
 KW Human; kinase; PKIN-17; cancer; leukaemia; adenocarcinoma; osteoporosis;  
 KW immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;  
 KW Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;  
 KW allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;  
 KW autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;  
 KW Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;  
 KW rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;  
 KW hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;  
 KW cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;  
 KW congestive heart failure; ischemic heart disease; lung tumour; gout;  
 KW fatty liver; Niemann-Pick's disease; gene therapy.  
 KW  
 XX  
 OS Homo sapiens.  
 FH Key location/Qualifiers  
 FT Domain 50...301  
 FT Domain /note="Eukaryotic protein kinase domain"  
 FT Domain 51...292  
 FT Domain /label="Protein\_kinase\_domain"  
 FT Domain 52...292  
 FT Domain /label="Protein\_kinase\_domain"  
 FT Domain 71...292  
 FT Domain /label="Protein\_kinase\_domain"  
 XX W0200196547-A2.  
 XX  
 XX 20-DEC-2001.  
 PF 14-JUN-2001; 2001WO-US19444.  
 XX  
 PR 15-JUN-2000; 2000US-212073P.  
 PR 23-JUN-2000; 2000US-213467P.  
 PR 30-JUN-2000; 2000US-215651P.  
 PR 07-JUL-2000; 2000US-216605P.  
 PR 13-JUL-2000; 2000US-218372P.  
 PR 25-AUG-2000; 2000US-228056P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;  
 PI Gandhi AR, Tribouley CM, Walla NK, Yao MG, Lu DM, Greenwald SR;  
 PI Bankumart J, Griffin JA, Kearney L, Butford N, Nguyen DB, Tang YT;  
 PI Baughn MR, He A, Thornton M, Hafalla A, Patterson C, Gururajan R;  
 PI Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JJ, Ding L;  
 PI Grether M, Elliott VS, Thangaveilu K, Batra S, Ison CH;  
 XX  
 DR WPI: 2002-090207/12.  
 DR N-PDB: AAD26464.  
 XX  
 PT New polypeptides, useful for diagnosing, treating or preventing  
 PT disorders of growth and development, cardiovascular and lipid, and  
 PT diseases such as cancer, comprise human kinase polypeptides -



Seq	Sequence	674 AA:	68.0%;	Score 2794;	DB 23;	Length 674;
Query Match						
Best Local Similarity			73.1%;	Pred. No. 4.5e-180;		
Matches 549;			Conservative 55;	Mismatches 49;	Indels 98;	Gaps
OY	1	MSGAGKEGGGSPAHNLPHRPHRPHQHAQYVPRYLEKTLGGGQGVNKLGYHCLTGQAV	60			
Db	1	MSSTRKDDGA-----OHAYGVPRLEKTLGGQGVNKLGYHCLTGQAV	45			
OY	61	AIKIVNREKLSSEVLMKVEREIAIILKLEHPVLMKHIDVENKKLYLVLEHVGSELFD	120			
Db	46	AIKIVNREKLSSEVLMKVEREIAIILKLEHPVLMKHIDVENKKLYLVLEHVGSELFD	105			
OY	121	YLVKGRLLTPKFAKRFPROIVSALDFCHSHYSICHDRLPENILLDEKNNIRIADFGMASL	180			
Db	106	YLVKGRLLTPKFAKRFPROIVSALDFCHSHYSICHDRLPENILLDEKNNIRIADFGMASL	165			
OY	181	QVGDLSLLETSCGSPHYACPEVTKGKRYGGRADMMSCGYILFALLVGALEPFDDMLROL	240			
Db	166	QVGDLSLLETSCGSPHYACPEVTKGKRYGGRADMMSCGYILFALLVGALEPFDDMLROL	225			
OY	241	EKKRGVFMHPFPIPPDCOSLLRGMIIEVEPEKRLSLEOIKHPWYLGKHEHEDPCTEPAP	300			
Db	226	EKKRGVFMHPFPIPPDCOSLLRGMIIEVEPEKRLSLEOIKHPWYLGKHEHEDPCTEPAP	284			
OY	301	GRRVARSLLPSNGELDPVLESMAISGCFRDERLHRELRLSEENQEKMIYLLDRKER	360			
Db	285	RKVOIRSLPSLEDIDPVLDSMHSIGCFRDRNKLLODLSSEENQEKMIYLLDRKER	343			
OY	361	YPSCDOLPPRNDVDPFRKRYDSPMLSRHGKRRRPERKSMEVLSTTDAGCGSSPVPTARRA	420			
Db	344	YPSOEDQLPPRNEIDDPFRKRYDSPMLSRHGKRRRPERKSMEVLSTTDAGCGSSPVPTARRA	400			
OY	421	LEMAHSORSRSVSASGATSSPLSPSPRSVPFSFSPSPGAGDEARGGSPSTKQTLPS	480			
Db	401	LEMAHSORSRSVSASGATSSPLSPSPRSVPFSFSPSPGAGDEARGGSPSTKQTLPS	429			
OY	481	RGRPGGAGEOPPPPSABASTPLPGPPGSPRSSGGTPLHSLPTTPRASPTGTTPPSP	540			
Db	430	-----YTPHSPSPGSPLPFPKG-----TPVHTPKESPAGPNNPFPSP	468			
OY	541	GGGVGAGAMRRLNLSIRNSFLGSPFRHRRKQVPALEKMSLTPSSPELAKRSWFGNFI	600			
Db	469	--SVGGVWRARLNSIKKSLFSGSPFRHRRKQVPALEKMSLTPSSPELAKRSWFGNFI	526			
OY	601	SLDKEQGFVLKDPLESLIKADIVHARLSPISLSHVSLSOTSPFAEKKAGSGSPVPOKP	660			
Db	527	SLDKEQGFVLKDPLESLIKADIVHARLSPISLSHVSLSOTSPFAEKKAGSGSPVPOKP	586			
OY	661	VRFQVDISSSEGPSPRRDSSGGGIGIVSYFTLLISGSPRRFRKRVETIQALLSTHDOP	720			
Db	587	VKFOVDIYTEGGEQKQK-----NGIYSVFTLLISGSPRRFRKRVETIQALLSTHDOP	640			
OY	721	SVQALADEKNGAQTRPAGAPRSLQPPGPR	751			
Db	641	AAQHLS-----EPPPPAP	653			
RESULT 4						
AAE21723						
ID	AAE21723	standard; Protein; 664 AA.				
XX	AAE21723;					
XX	AC					
XX	DE	16-JUL-2002 (first entry)				
XX	XX					
XX	XX	Human PKIN-18 protein.				
XX	XX	Human; kinase; enzyme; PKIN-18 protein; immune system disorder; anaemia;				
XX	XX	acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;				
XX	XX	asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;				

KW	AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;
KM	leukemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;
KV	Down's syndrome; gene therapy; protein therapy; cytoskeletal.
XX	
OS	Homo sapiens.
FH	Key
FT	Domain
FT	/note= "Eukaryotic protein kinase domain"
FT	15..266
FT	16..257
FT	/note= "Protein kinase domain"
FT	17..257
FT	/note= "Protein kinase domain"
FT	36..256
FT	/note= "Protein kinase domain"
PX	
PN	WO200218557-A2.
XX	
PD	07-MAR-2002.
PF	31-AUG-2001; 2001WO-US27219.
PR	31-AUG-2000; 2000US-229873P.
PR	08-SEP-2000; 2000US-231357P.
PR	14-SEP-2000; 2000US-232654P.
PR	22-SEP-2000; 2000US-234902P.
PR	29-SEP-2000; 2000US-236499P.
PR	06-OCT-2000; 2000US-238389P.
PR	13-OCT-2000; 2000US-240542P.
PA	(INCY-) INCYTE GENOMICS INC.
PI	Bangman O, Nguyen DB, Wallia NK, Hafalia AJA, Yao MG, Gandhi AR;
PI	Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Trimbouley CM;
PI	Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YT;
PI	Azizmal Y, Burrill JD, Marcus GA, Zingler KA, Lu DAM, Lai PG;
PI	Ramkumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;
PI	Burford N;
DR	WPI: 2002-329769/36.
DR	N-PSDB: AAD34315.
PT	New human kinases, useful for diagnosing, treating or preventing immune
PT	system disorders (e.g. Crohn's disease), neurological disorders (e.g.
PT	epilepsy), or cell proliferative disorders (e.g. cancers such as
PT	leukemia or lymphoma)
PS	
XX	Claim 73; Page 181-183; 218pp; English.
CC	The present invention relates to human kinases (PKIN) and polynucleotides
CC	encoding such proteins. PKIN sequences of the invention are useful for
CC	diagnosing, treating or preventing disorders associated with aberrant
CC	expression of PKIN, particularly immune system disorders (e.g. acquired
CC	immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
CC	asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
CC	Tooth disease or seizures); cell proliferative disorders (e.g. cancers
CC	such as adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma),
CC	and developmental disorders (e.g. Down's syndrome). They are also used
CC	in gene therapy and protein therapy. The present sequence is human
CC	PKIN-18 protein.
SQ	Sequence 664 AA;
Query Match	67.4%; Score 2767.5; DB 23; Length 664;
Best Local Similarity	76.9%; Pred. No. 2.7e-178;
Matches 539; Conservative	52; Mismatches 45; Indels 65; Gaps 7;
D0	11 YGGRLEKTKGGTGLVKLGVCITCGKVAIKIYNREKLSSESYLMKVERETAIKLKE 89
D0	YGGRRLEKTKGGTGLVKLGVCITCGKVAIKIYNREKLSSESYLMKVERETAIKLKE 70
D0	HPVLKLDHYENKKRYLYLVIEHVSGGELFDYLAVKKRLLTPRKARFFQIVSALDFCHS 149

```

Db 71 HPHVILKADHYENKYYLYLVLEHVSQGELEFDYLVKKGRLLTPKARFFROIISALDFCHS 130
Oy 150 YSICHRDLKPEMLLDEKNNIRIADFGMASLQYDLSLETSQSPHYACPEVIGKEKYD 209
Db 131 HSHCHDLKPEMLLDEKNNIRIADFGMASLQYDLSLETSQSPHYACPEVIGKEKYD 190
Oy 210 RRADWMSGVILLFALLVGLPDDNNLRQLLEKVKGVFMHPIPPDCOSLLRGMIEVE 269
Db 191 RRADWMSGVILLFALLVGLPDDNNLRQLLEKVKGVFMHPIPPDCOSLLRGMIEVE 250
Oy 270 PERKRLSDIOKHHPWLVGKQKHEPDLCPAPGRVAMRSLSNGELDDPVLESMASIGCF 329
Db 251 AARLLLEHIOKHWYIGGNEPEP-EQPIP-RKVOIRSLLEDIDDPVLESMASIGCF 308
Oy 330 RDERLHRELSEENQEKMIYLLDRKERYPSCEDQDLPNNVDYPPKRVDSPLMSR 389
Db 309 RDRNKLQDLSSEENQEKMIYLLDRKERYPSCEDQDLPNNVDYPPKRVDSPLMSR 368
Oy 390 HGRKRPERSMEVLTITDAGGSGPVPTRRALMAQHSQSRSVSGASTGLSSPLSSPR 449
Db 369 HGRKRPERSMEVLTITDAGGSGPVPTRRALMAQHSQSRSVSGASTGLSSPLSSPR 425
Oy 450 SPVFSRSPERGADDEARGGSPSTKQTLPSRGRPGGAGEOPPPPARSGTLPBPGPSP 509
Db 426 -----VTPHPSRSPPLTPFKG-- 442
Oy 510 RSSGCTPLHPLTPRASPPTGPTGTPPPSPGGGAGVGAAMRSRLNSIRNGLSPRPHR 569
Db 443 -----TPVHTPKRESPTAGTNPPTPPSP--SVGGVPPRARLNSIKNSPLSGPRPHR 491
Oy 570 KMOVPAEAKMSITPSSPEPLAKRSWFGNLTSLDKKQITLVLDKPLSSIKADIVAF 629
Db 492 KLOVPPPEKMSNTLPSSPEPLAKRSWFGNLTSLDKKQITLVLDKPLSSIKADIVAF 551
Oy 630 STPLSHSVYLSQTSFPAEYKASGSPVFOKPVPRQVODISSEGEPEPBRDGSQGGT 689
Db 552 STPLSHSVYLSQTSFPAEYKASGSPVFOKPVPRQVODISSEGEPEPBRDGSQGGT 605
Oy 690 VPEFTLISGPRRRKRVETIOAQLSLTHDQPSVALDEKN 730
Db 606 VPEFTLISGPRRRKRVETIOAQLSLTHDQPSVALDEKN 646

RESULT 5
ABB62061
ID ABB62061 standard; Protein; 851 AA.
XX
AC ABB62061:
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 12975.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
  pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN MO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001MO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PMD, Myers EW,
DR WPI; 2001-656860/75.
DR N-PSDB; ABL06164.

```

```

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
PS Disclosure; SEQ ID NO 12975; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB157737-AB170722).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 851 AA:
SQ

```

```

Query Match 47.5%; Score 1952; DB 22; Length 851;
Best Local Similarity 54.9%; Pred. No. 3.6e-123;
Matches 421; Conservative 96; Mismatches 138; Indels 112; Gaps 24;

```

```

Oy 26 QHAQVGRYLEKTLGKQGTGLVLCVHTTGQVVAIKYVREKLSVLMKVEREAIL 85
Db 10 ENCOFVGRYLEKTLGKQGTGLVLCVHTTGQVVAIKYVREKLSVLMKVEREAIL 69
Oy 86 KLIHPHYLKHADHYENKYYLYLVLEHVSQGELEFDYLVKKGRLLTPKARFFROIISALD 145
Db 70 KLIHPHYLKHADHYENKYYLYLVLEHVSQGELEFDYLVKKGRLLTPKARFFROIISALD 129
Oy 146 FCHSYICHRDLKPEMLLDEKNNIRIADFGMASLQYDLSLETSQSPHYACPEVIGKE 205
Db 130 FCHSHICHRLKPEMLLDEKNNIRIADFGMASLQYDLSLETSQSPHYACPEVIGKE 189
Oy 206 KYDGRADWMSGVILLFALLVGLPDDNNLRQLLEKVKGVFMHPIPPDCOSLLRGM 265
Db 190 KYDGRADWMSGVILLFALLVGLPDDNNLRQLLEKVKGVFMHPIPPDCOSLLRGM 249
Oy 266 IEVEPERKLSLEIOKHWPY-LGGKHEPD--PCLEPAPGRVAMRSLSNGELDDPVLE 321
Db 250 IEVNDPRRLTAEINRHWYVWAGGKGELELEIPMEV---VQTHVPTATADPDVLA 304
Oy 322 SMASIGCFRDERRLHRELSEENQEKMIYLLDRKERYPSCEDQD--LPPND--V 375
Db 305 AICSLGCFKEKELIQELSSHNTKTYFLLERKRRRPALEDDEDLAQSRELDAY 364
Oy 376 DPPKRVDSPLMSRHRKRPERSKSEVLTIDAGGSGSPV-PTRRALMAQHS---QRS 430
Db 365 DPPKRVDSPLMSRHRKRPERSKSEVLTIDAGGSGSPV-PTRRALMAQHS---QRS 413
Oy 431 RSVASGASTGLSSPLSSP--RSPVFS-----FSPPEGA-----GDEARG--GS 470
Db 414 RSPVTVTSVSVSSYHSPTRCNSPMSAQAQAMASRSPPAAGRHSHTYGDRODSGHS 473
Oy 471 PTKTQTLPS-----RGRPGGAGEQPPPARSTPLFG-PGG 507
Db 474 SVSRTPSHSQSIGDYVYVVEPRTERDSLRQERGG-----SPRDGDGIGPG 525
Oy 508 SP--RSSGCTPLHPLH-----TPRASP--TGTPGTTPPSPGGAGGAAW 549
Db 526 SPGNSGSGTASPSVHRANSPTIATIVNPGSPMANNSSPGM--PGSPONTPGQILW 583
Oy 550 RSRLLSIRNSFLGSPRFRKMQVPTAEEMSLTPSSPELAKRSWFGNLTSLDKKQIT 609
Db 584 KTRLTNINNSFLGSPRFRKMQV-SADEV-HLTPESSPELAKRSWFGNLTSLDKKQIT 641
Oy 610 LVLDKPLSSIKADIVAFPLSTPSLSVLSQTSFPAEYKASG-GPSVFOKRVPRQVDIS 668
Db 642 ILVKGKPIATYKAHLIHAFLSMALSHSVSPTSPRVEYKRNNGPVMVFORHVKFOVDIS 701

```

QY 669 SSEGPPSPRRDGGGGGIVSTFTLLSGPSRRKRVETIOAQLLS 715  
 DB 702 AI-----CKGDIADMLFALFTFTLLSGNTRRRRCHEHQVCS 741

## RESULT 6

ID ABB11052 standard; peptide; 301 AA.

AC ABB11052;

DT 11-JAN-2002 (first entry)

DE Human HRPDK-1 homologue, SEQ ID NO:1422.

Human; cytokine; cell proliferation; cell differentiation; growth factor;  
 haematopoiesis regulation; tissue growth; immunomodulator; activin;  
 inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
 myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 chronic inflammatory condition; proliferative retinopathy;  
 atherosclerosis; coronary heart disease; arterial ischaemia;  
 bone disorder; osteoporosis; vascular growth disorder;  
 tissue regeneration; wound healing; infection; immune disorder;  
 cell culture; drug screening; gene therapy; antiinflammatory;  
 antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
 cytoskeletal; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
 antifungal; vulnery; antitumor.

OS Homo sapiens.

PN WO200157188-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US03800.

PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PA (HXSE-) HXSEQ INC.

PI Tang YT, Liu C, Dermanac RT;

XX WPI: 2001-457740/49.

DR N-PSDB; ABA08296.

XX Human proteins and DNA encoding sequences useful for preventing,

PT treating or ameliorating a medical condition in a mammalian subject

PT e.g. arthritis and cancer.

XX Claim 20; Page 140; 1963pp; English.

Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
 sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
 invention also relates to vectors and recombinant host cells comprising a  
 nucleotide of the invention, methods of producing the novel polypeptides,  
 antibodies against the polypeptides, methods of detecting the nucleotides  
 or polypeptides in a sample, and methods of identifying compounds which  
 bind to polypeptides of the invention. Although novel, many of the  
 polypeptides of the invention have homology to known proteins, thereby  
 giving an insight into their probable biological activities, and hence  
 potential therapeutic applications. The polypeptides of the invention may  
 have various activities, including cytokine, cell proliferation or cell  
 differentiation activities; stem cell growth factor activity;  
 haematopoiesis regulatory activity; tissue growth activity;  
 immunomodulatory activity; activin- or inhibin-related activities;  
 chemotactic or chemokinetic activities; haemostatic, thrombotic or  
 thrombolytic activities; receptor or ligand activities; or may be  
 involved in oncogenesis, cancer cell proliferation or metastasis.  
 Depending on their biological activities, polypeptides and nucleotides of  
 the invention are useful for preventing, treating or ameliorating medical  
 conditions, e.g., by protein or gene therapy. Such conditions include

CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
 CC vascular growth. Polypeptides involved with tissue regeneration and  
 CC repair (or nucleic acids encoding them) may be used to promote wound  
 CC healing (e.g., of burns, incisions and ulcers), while those with  
 CC immunomodulatory activities may be used in the treatment of viral,  
 CC bacterial and fungal infections in addition to immune disorders.  
 CC Polypeptides with growth factor activity may be used in cell cultures to  
 CC promote cell growth. For example, such polypeptides may be used to  
 CC manipulate stem cells in culture to give rise to neuroepithelial cells  
 CC that can be used to augment or replace cells damaged by illness,  
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
 CC may also be used in the diagnosis of the above conditions, and in drug  
 CC screening techniques. The present sequence represents a novel human  
 CC polypeptide of the invention.

CC Sequence 301 AA;

CC Query Match 32.7%; Score 1341.5; DB 22; Length 301;

CC Best Local Similarity 87.2%; Pred. No. 1.5e-82;

CC Matches 259; Conservative 1; Mismatches 2; Indels 35; Gaps 1;

QY 27 HAQVGPRLKTLKGGTGLVKGVCITGOKVAIKIYRKLSSVLMKVERETAILK 86

DB 1 HAQVGPRLKTLKGGTGLVKGVCITGOKVAIKIYRKLSSVLMKVERETAILK 60

QY 87 LIEHPVILKLVYENKRYL-----LYLVE 111

DB 61 LIEHPVILKLVYENKRYLPPDELTSGPSMLAQVSPHGKLSARRSMDLLSGEPYLVLE 120

QY 112 HVSGETLDYLVKRGRLPKEARKPPROIVSALDPCHSISICHRDLKPEMLIDKNNIR 171

DB 121 HVSGETLDYLVKRGRLPKEARKPPROIVSALDPCHSISICHRDLKPEMLIDKNNIR 180

QY 172 IADFGMASLQVDSLSLETSCGSPHYACPEVIRGEKYGRRADMSCGVILFALLVAGLPE 231

DB 181 IADFGMASLQVDSLSLETSCGSPHYACPEVIRGEKYGRRADMSCGVILFALLVAGLPE 240

QY 232 DDDNRLQLLEKVRKGVFNHPHPIPPDCOSLNGMFEVEPEKRLSLQIOKHWPYLCG 288

DB 241 DDDNRLQLLEKVRKGVFNHPHPIPPDCOSLNGMFEVEPEKRLSLQIOKHWPYLCG 297

## RESULT 7

ID ABB11680 standard; peptide; 204 AA.

AC ABB11680;

DT 11-JAN-2002 (first entry)

DE Human Ser/Thr kinase homologue, SEQ ID NO:2050.

Human; cytokine; cell proliferation; cell differentiation; growth factor;  
 haematopoiesis regulation; tissue growth; immunomodulator; activin;  
 inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
 myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 chronic inflammatory condition; proliferative retinopathy;  
 atherosclerosis; coronary heart disease; arterial ischaemia;  
 bone disorder; osteoporosis; vascular growth disorder;  
 tissue regeneration; wound healing; infection; immune disorder;  
 cell culture; drug screening; gene therapy; antiinflammatory;  
 antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
 cytoskeletal; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
 antifungal; vulnery; antitumor.

OS Homo sapiens.

PN WO200157188-A2.

PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US03800.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI: 2001-457740/49.  
 N-PSDB: ABA08924.  
 XX  
 PT Human proteins and DNA encoding sequences useful for preventing,  
 PT treating or ameliorating a medical condition in a mammalian subject  
 PT e.g. arthritis and cancer -  
 XX  
 PS Claim 20: Page 227; 1963pp; English.

Sequences ABA10981-AB12330 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoietic regulatory activity; tissue growth factor activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, CC autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a novel human polypeptide of the invention.

Sequence 204 AA;

Query Match 25.4%; Score 1043; DB 22; Length 204;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-62;  
 Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 PTAEMSLTPRESSPELAKRSMFNFISLDEEQLFLVLRKPLSSIRADIYHAFSLIPS 633  
 DB 1 PTAEMSLTPRESSPELAKRSMFNFISLDEEQLFLVLRKPLSSIRADIYHAFSLIPS 60  
 QY 634 LSHSVLSQTSFRAEYKASGSPVFOKPVRFQVNDISSSGEPSPRRDGGGGGIVSVTFT 693  
 DB 61 LSHSVLSQTSFRAEYKASGSPVFOKPVRFQVNDISSSGEPSPRRDGGGGGIVSVTFT 120  
 QY 694 LISGSRFRKRVVEIIQALSTHDQPSVQALADEKNGAQRFPAPPRSLQPPGRDP 753

DB 121 LISGSRFRKRVVEIIQALSTHDQPSVQALADEKNGAQRFPAPPRSLQPPGRDP 180  
 QY 754 ELSSPRRGPPKDKKLATNGTPL 777  
 DB 181 ELSSPRRGPPKDKKLATNGTPL 204

RESULT 8  
 AAG70854  
 ID AAG70854 standard; Protein; 1349 AA.  
 XX  
 AC AAG70854;  
 XX  
 DT 27-JUL-2001 (first entry)  
 XX  
 DE C albicans apoptosis associated protein #34.  
 XX  
 KW Yeast; fungus; apoptosis; infection; proliferative disease;  
 KW vaccine; autoimmune disease; Ischaemia; neurodegeneration.  
 XX

XX Candida albicans.  
 XX WO200102550-A2.  
 XX 11-JAN-2001.  
 XX 03-JUL-2000; 2000WO-BE00077.  
 XX 01-JUL-1999; 99EP-0870141.  
 XX (JANC) JANSEN PHARM NV.  
 XX Contreras RH, De Backer MD, Luyten WHML, Malcorps IKL;  
 PI Nelissen BJM, Reekmans RJ;  
 XX WPI: 2001-367042/38.  
 DR N-PSDB: AAH29890.

PT Yeast and fungal nucleic acids encoding proteins involved in a pathway  
 PT leading to programmed cell death, useful for treating proliferative  
 PT disorders, yeast and fungal infections, or for preventing apoptosis in  
 PT certain diseases -  
 XX  
 XX Claim 24; Fig 2; 218pp; English.

CC The present invention provides the protein and coding sequences of a  
 CC number of apoptosis associated proteins from the yeast *Saccharomyces*  
 CC cerevisiae and the fungus *Candida albicans*. These can be used to identify  
 CC treatments for fungal and yeast infections, for proliferative diseases  
 CC and for apoptosis related diseases such as autoimmune diseases, ischaemia  
 CC and neurodegeneration. The present sequence is one of the C. albicans  
 CC proteins of the invention.  
 CC  
 XX

Sequence 1349 AA;

Query Match 21.0%; Score 862; DB 22; Length 1349;  
 Best Local Similarity 30.3%; Pred. No. 2.1e-49;  
 Matches 254; Conservative 112; Mismatches 321; Indels 150; Gaps 22;

QY 20 PHHPP-----OHAQVPPYRLKTKTGKGTGLVLYGVHCTGKVAIKIVNR 67  
 DB 2 PHSRQPSISSISMSOSNNHHPKIGPWKILGKATGATGATVLAHQTTGGAAKVVSK 61  
 QY 68 ELSSSVLMK-----VERIATILKLIENHPVLRKLDVYENKRYLYLVRVSGGELE 119  
 DB 62 SELQDEETKKNDDGLPYGIEREITIMKLTTHPNVLRXYVWETSRALYLVLEVEGGELE 121  
 QY 120 DYLVRKGLTPKAPKFFQIYSALDFCHSYISICHRDLKPNMLLDEKNNINIAIDGMS 179  
 DB 122 DILVRGRLPEVEAKIKYFQIILGTAYCHALGICHRDLKPNMLLDSQNLVLRADGMA 181  
 QY 180 LGVDSLLETSCGSPHYACPEYIKGEKYDGRADMMSCGVIIFALLVGLALPPDDNRL 239

```

DB 182 LBSNGKLLFTSCGSPHIAPEIVSGLEKYGASDVMSCGYLLFALLGRLPEDEENINL 241
QY 240 LKVKRGVEHMP-HFIPPCOSILRGMIIEVEPEKRLLEOIKRHPW-----YLG 288
DB 242 LKVGAGNFEMPYDEVSRAROLDIARLEVDPMRISTEKILRPLDKYPMSEDLISE 301
QY 289 KHEPDCLEPARBARRAMSLPSNGELDPVLESMSLGCFFRDRERLHRELSSENOEK 348
DB 302 KSLPHP-----HTGKSLGIVRNIDKILSLTLIMDRPEEIVDLCKGDSNEK 353
QY 349 MIYLLDLKERKPYSCEDDLPVRNDVDPKRYDSMLSRHGKR-----RPEKSMYVL 403
DB 354 TYVALIM-----KKHMDNN-----TNNSPKSKTSFNKKVYRSGSKYILNGTPRRKRASHI 406
QY 404 ST-----TDAGGGGSPVPTRRALEMAQHSORNSVSGASTGLSSPLSP-RRPV- 452
DB 407 SVSRPTSFQYKSNPGAGA-----TANRNSVARHSVASSANNSPKRSPYKSRYSRSPYK 462
QY 453 -----FSESPGAGDEAGKGGG-----PSTKQTL 478
DB 463 SPSKRYSYNQSPKSPYGRNSQROFENEPKAKPRNIYNEIYDAQSNFSLPPLPSL 522
QY 479 PSRGPRGGAGGPPPPSARSTPLPGPPSGRSGGTPHSPHTPASPCTGTCTPPP 538
DB 523 PSKDSR--YIMDEPNQPOLQAPALSOVPENP-----IYDESPDL 559
QY 539 SPGGVGGAGMRSRLNSIRNSFLSGSPRFHRRK---MQVPAEEMSLTPESPPELAKRW 595
DB 560 MOSAKISSGRNSIIGKNNNSNSNKRMSKRKISRASMTGLKRNSTMLLTSLYATLGG 619
QY 596 FGNFSLKREQ-----IFVLKDKPLSSIKADIVHAFSLPSLSH-----VLSQTSF 644
DB 620 DDDMYMDQIKRTSATFALCDKIFNOEDYDEEDQLVDPEEKAKVEYERLMELERKKH 679
QY 645 RAERYASGGSPYFQKPVPRQV--DISSEGEPEPSPRRDSGGGIGTSVFTLLSGPSRRF 702
DB 680 EAELEKAR--RELEKKRKRKRSLISLKKLSTIVYNDADPNNSGEQIVDEGIKOPRRQS 736
QY 703 KRVVETIOQLSTHDQSVQALADEKKAQCTPPACAPPSRLIOPPGRRPDELSSSP 759
DB 737 KNL--ALRALSEGNSHASELLEDEVENILKRSASQP-----VPRKRQPTVLTTRP 785

```

RESULT 9  
AAE21712  
ID AAE21712 standard; Protein: 1369 AA.  
XX  
AC AAE21712;  
XX  
DT 16-JUL-2002 (first entry)  
XX  
DE Human PKIN-7 protein.  
XX  
KW Human; kinase; enzyme; PKIN-7 protein; immune system disorder; anaemia;  
KW acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;  
KW asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;  
KW AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;  
KW leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;  
KW Down's syndrome; gene therapy; protein therapy; cyostatic.  
OS Homo sapiens.  
XX  
XX  
FH Key location/Qualifiers  
FT Peptide 1..14  
FT /label= signal\_peptide  
FT /label= 15..1369  
FT /label= Human\_mature\_PKIN-7\_protein  
FT /label= 66..317  
FT /note= "Eukaryotic protein kinase domain"  
FT /note= 67..308  
FT /note= "Protein kinase domain"  
FT Domain 70..305

```

FT /note= "Protein kinase domain"
PN W020218557-A2.
XX 07-MAR-2002.
XX 31-AUG-2001; 2001WO-US27219.
XX 31-AUG-2000; 2000US-229873P.
XX 08-SEP-2000; 2000US-231357P.
XX 14-SEP-2000; 2000US-232554P.
XX 22-SEP-2000; 2000US-234902P.
XX 29-SEP-2000; 2000US-236499P.
XX 06-OCT-2000; 2000US-238389P.
XX 13-OCT-2000; 2000US-240542P.
XX (INCYTE GENOMICS INC.
XX
XX Bhandman O, Nguyen DB, Walla NK, Hafalia AJA, Yao MG, Gandhi AR;
XX Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;
XX Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YT;
XX Azimzai Y, Burdill JD, Marcus GA, Zingler KA, Lu DAM, Lai PG;
XX Rankumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;
XX Burford N;
XX WPI; 2002-329769/36.
XX N-PSDB; AAD34304.
XX
XX New human kinases, useful for diagnosing, treating or preventing immune
XX system disorders (e.g. Crohn's disease), neurological disorders (e.g.
XX leukaemia) or cell proliferative disorders (e.g. cancers such as
XX lymphoma)
XX Claim 62; Page 152-156; 218pp; English.
XX
XX The present invention relates to human kinases (PKIN) and polynucleotides
XX encoding such proteins. PKIN sequences of the invention are useful for
XX diagnosing, treating or preventing disorders associated with aberrant
XX expression of PKIN, particularly immune system disorders (e.g. acquired
XX immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
XX anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
XX Tooth disease or seizures), cell proliferative disorders (e.g. cancers
XX such as adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma),
XX and developmental disorders (e.g. Down's syndrome). They are also used
XX in gene therapy and protein therapy. The present sequence is human
XX PKIN-7 protein.
XX
XX Sequence 1369 AA:
SQ

```

Query Match 20.7%; Score 851; DB 23; Length 1369;  
Best Local Similarity 29.4%; Pred. NO. 1.2e-48;  
Matches 258; Conservative 124; Mismatches 291; Indels 204; Gaps 30;

```

QY 2 SSGAKEGGGSPAHLPHP-----HHPHQQAQ-----YGPYRLKLT 39
DB 12 AAGGTGGAGAGRLLPPAPGSPAPAAVSPAAQPPRPAPASGMPARGYEIDRT 71
QY 40 LKQGTGLVKGVCITGQKVAIKIVNREKLSSEYIMKVERITALIKLEPHVHLKHY 99
DB 72 IGRGNFVVRKATHLVTKAKAIKIDTQDDENLKIFREYQIMKMLCHPHTRLVY 131
QY 100 YENKKYLYLVLEHVSGLFEDYLVKKGRLTREKARFFQIYVALDFCHSYICHRDLKP 159
DB 132 METERMITYLVTEYASGGEIFDHLVAHGMAEAKRFRKQIYVYFCHCRNIVARDLKA 191
QY 160 ENLLDEKNNIRIDFGASIQVDSILLETSCGSHYACPEYIKCEKDGRRADMSGCV 219
DB 192 ENLLDNLNMLKIDAFSNTFTPGQLTKWCSPPYAPAEFEKKEYDGPVDTWISLGV 251
QY 220 ILFALLVGALEPDDNIRQLLEKVRGVGFHFPDCCOSILRGMIIEVEPEKRLISQI 279
DB 252 VLVYVCGALPFDGSLTQNLARAVLSGKFRIFPFMSTECHEILIRMLVLDPKRLSMQI 311

```

```

QY 280 QKHPVTLGKHEPDCLEPAPGRVA-MRSLPSNGELDP---DYLESMAISGCFDRERL 335
DB 312 CKRKMWKLQ---DADPNPD---RLIACQOOLKEBROVDPLNEDVLAMEDMGL--DKBQT 363
QY 336 HRLRSEENQEMKYYLLLDKREYPSCEQDDP--PR-----NDV 375
DB 364 LQSLRSDAYHYSAIYSLCDKRRKHTLRGALPSMPRALAFQAPVIOAEQAGTAMNI 423
QY 376 DPKRKYSDPMLSRHGRKRRERKSM---VLST-TDAGGSGPVPTRRALEMAQS--- 427
DB 424 SVQVQVLIN-----DENQIVEPDGTLNLDSDGESESPALVYLSMRRTVGVA 473
QY 428 -QSRSTVSGASTGLSSPSLSPSPVSPFSPPEAGDEARGGSTKTQTLPSGPRGG 486
DB 474 DPKTEWEDLQKLLPGPPGVNPQAPFLQVAP-----VNFNMHNLPMQNLPT 521
QY 487 GAGEOP-----PPSARSTPLPGPPGSPRSRG--PLHSP--LHTPPA-SPTGT----- 531
DB 522 GQLEKESQLQPTTLLQNGMPLGRASDGANTQLHAQOLKRPSPSLVTMTAV 581
QY 532 PGTPP--PPSPGGVGAAMRSRLNSTRNSFLGSPRHRKMOVPTAEEMSSLTPESSP 588
DB 582 PATVPDESSDGPDEAVORRYLANRSKR-----HTLAMTPRAE---IPDLQR 629
QY 589 ELAK-----RSWGNFSLDKEDIFLYLAKPLSSIK-ADIVH---AFSLPSLSHV 638
DB 630 QLGQOQPFRRSHVMPHLV-----PDQHRSTYKDSNTLHLPTERFSPRRSDIA 677
QY 639 LSGTSPFAEKKASGGS-----VFQKPVRFQVDISSEGEPEPRRDGSGGGIX 688
DB 678 ASIQAFRAHLEKMGNSIKOLOECQOLKMGQID----- 715
QY 689 SVFTLLISGSRREKRVETIQALLSTHDQPSVQALADENKGAQTRPAGAP----- 741
DB 716 -----ERTLEKYOQHMLYQOEQHIIQOQIODSICPPQSPPLQACEN 761
QY 742 -----RSLQ-----PPGRDPPELSSSPRGRPP 764
DB 762 QPALTRHQIQLRIQPSPPNHPNHLFRQPSNSPP 798

RESULT 10
AAM90879
ID AAM90879 standard; Protein; 823 AA.
XX
AC AAM90879;
XX
DT 07-JUL-2000 (first entry)
XX
DE Human keratinocyte derived pke#122 protein #2.
XX
KW Keratinocyte; regulatory protein; human; pke#122; antiproliferative;
KW kinase; signal transduction; desmosome; dermatological condition;
KW pemphigus vulgaris; psoriasis vulgaris; epidermis; cosmetic.
XX
OS Homo sapiens.
XX
PN WO200017232-A2.
XX
PD 30-MAR-2000.
XX
PE 06-SEP-1999; 99WO-DE02865.
XX
PR 19-SEP-1998; 98DE-1042863.
XX
PA (KRAM/) KRAMER M.
XX
PI Kramer M, Bechtel M, Reinartz J, Schaefer B, Wallich R;
XX
DR WPI; 2000-283542/24.
XX
DR N-PSDB; AAX82952.
XX
PT New regulatory polypeptide from keratinocytes, useful for diagnosis and

```

```

PT treatment of dermatological disorders, also in cosmetics
XX
XX Claim 1; Page 53; 55pp; German.
PS
XX
CC This invention describes a novel human regulatory polypeptide designated
CC pke#122 (I), or its variants from human keratinocytes, which in activated
CC keratinocytes is the same as, or similar to, the most highly expressed
CC protein and has antiproliferative activity. (I) is a kinase involved in
CC signal transduction and is conjectured to be involved in cell-cell or
CC cell-matrix interactions and/or formation of (heml)desmosomes. Modulation
CC of (I) will affect proliferation and differentiation of keratinocytes.
CC (I) is upregulated in some dermatological conditions, e.g. pemphigus
CC vulgaris and psoriasis vulgaris. (I) is used to raise specific antibodies
CC (Ab) which are used for diagnosis and treatment of dermatological
CC disorders and for cosmetic treatment of the epidermis. Nucleic acid (II)
CC that encodes (I) is used; (I) to produce transgenic mammals (particularly
CC mice and rats); and (II) as source of sense and antisense
CC oligonucleotides for diagnosis and treatment of dermatological disorders
CC and for cosmetic treatment of the epidermis. (II), (I) and polypeptides
CC (Ia) produced by splice variants of (II) are used to identify specific
CC binding agents (potentially useful in medicine and cosmetics) that
CC modulate their function and/or expression, especially inhibitors and
CC activators. Cells and animals transformed with (II) are used to study
CC (I)-induced changes in cell morphology and basic functions. This
CC sequence represents the human keratinocyte protein pke#122 described in
CC the method of the invention.
XX
XX
SQ Sequence 823 AA;
XX
Query Match 20.5%; Score 840.5; DB 21; Length 823;
Best Local Similarity 29.9%; Pred. No. 3.2e-48;
Matches 256; Conservative 102; Mismatches 258; Indels 201; Gaps 27;
XX
QY 4 GAKGCGGSAVYHL-----PAPHPPQHAQVYGRYEKTLGKGTGLVKGVCII 55
DB 29 GRSGGGGAGAMAYINSEFASDPAQSGQOKSLRVGYDIERLTKGNFVYKLARHY 88
QY 56 TQGRVAIKIYNNREKLSSVLMKVEREIALKLEIHPYVLKLDHYENKYLVLVLEHVS 115
DB 89 TKTOVAIKIITDKTRLDSNNLEKIRYVOLMKLNMHPHILIKIYQWETMDMITYVTEFAKN 148
QY 116 GELFDYLVKKGRLPKRKRFRRQIVYALDFCHSYICHRDLKPNLLDEKNRIRIDE 175
DB 149 GEMFDYLTNSHLSSENDERKRFMOILSVEYCHDHVYHRDLKTNLLDGMNDIKLDF 208
QY 176 GMAISQYDLSLETSCGSPHYACEVIGEKYDGRADWMSGVTLFALLVGLPFPDDN 235
DB 209 GFGNFTYSGELSTWCSPPYAAAEVFEKYEGBQDLWISGLVLYLVGCSLFPDGN 268
QY 236 LRQLLEKVKRGVFMPIPIPDOSLRLGMEVEPERKLSLEQLOKHVWYLAGKEPPDC 295
DB 269 LPTLRQVYLERFRFIPFMSQDCESLIRMLVVPARITTAQIQHKKMM-----RAEPC 323
QY 296 LEPAPG-RVYAMRSLPSN-GEIADPVLESMAISGCFDRERLHRLRSEENQEMKIYTL 353
DB 324 L-PGPACPAPSAHYSTSLNGLDYDEQALGIMOTLGV--DRQKVESIQNSYHNFALITYL 380
QY 354 LLDKREYPSCEQDDLPRLNDVDPKRRKRVDPMLSRHG-KRRPERKSMELVSLTDAGGG 412
DB 381 LLERLKEKRYNQ-----CARPGPARQPRRSSDLSLEVQEGEL 419
QY 413 SPVPTRRALEMAQHSQKRSYSGASTG--LSSS-----PLSSRSPVFSFSP----- 457
DB 420 STDPEFRALLCPQPTLVQSVLAQEMDELSLQWLPFFPVDACSGCFRPRVSPSSL 479
QY 458 -EPGAGDEARGGSGPTSTQTLPSNGPRGGAGGQPPPSARSTPLPGP----- 505
DB 480 LDRTAISEPARQGPGLSEEDTQESL-PSSTGRHRTLAESTRLSLTAPCIYVSESTMS 538
QY 506 -----PGSPRSSGGTPLH-----SPLHTPRASP-TGTPGTTPPPSPGG 542
DB 539 PAEGTSSDSCLTFSASKSPAGISGTPTATGILGACSPVRL--ASPLFGQSATPTVLAAG 596

```

QY 543 GVGGA-----AMSRRLNLSINS--FLGSPRRH---RRKMOV 574  
 DB 597 GLGGAVLLPVSEFGRRASDTSLTGLKAFQQLRKTTRTGFLGLNKKIGLARQCVCP 656  
 QY 575 TAEEMSS-----LTPESPPEL-----AKRSMFGNFIIDKEQIFVLKDKPLSIKAD 623  
 DB 657 ASRASRGSLSPFARAPASPGIHLGAAGSRGEM-----688  
 QY 624 IYHAFSLPISLASHSVLSQTS- FRAEYKASGGPSVQKPRFOVDISSEGEPEPRR--- 679  
 DB 689 -----SLLEEVLEQQRLLQLQHHPAAAP-----GCSQAPQAPAPAPVY 726  
 QY 680 ---DSGGGGIYVYTF---LISGPRRRFRVYETIOAQLSTHDPQSVQALDEKNGA 732  
 DB 727 APCDGGAPAPLPTSLTSLTSGPLPPLPQLQTGASPVASAAQLDTH-----LHIGTGP 778  
 QY 733 QTRPAGAPPSRLQPPG 749  
 DB 779 TALPAVPPPRRLARLAPG 795

RESULT 11  
 AAG65764  
 ID AAG65764 standard; Protein; 783 AA.  
 AC AAG65764;  
 DT 07-JAN-2002 (first entry)  
 DE Human protein kinase polypeptide 3714.  
 DE  
 KM Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;  
 KM cytostatic; antiinflammatory; immunosuppressive; cardiant; hepatotrophic;  
 KM gene therapy.  
 OS Homo sapiens.  
 PN WO200173050-A2.  
 PD 04-OCT-2001.  
 PF 23-MAR-2001; 2001WO-US09483.  
 PR 24-MAR-2000; 2000US-191846P.  
 PA (MILL-) MILLENNIUM PHARM INC.  
 PI Meyers R;  
 DR WPI: 2001-611632/70.  
 DR N-PSDB; AAI66822, AAI66823.

New human protein kinase polypeptides, 3714, 16742, 23546 and 13887, useful in diagnosis of cancer or cellular proliferation or differentiation disorders and to screen for polypeptide modulators useful to treat such conditions -

Claim 4; Fig 1A-D; 169pp; English.

The invention provides novel human protein kinase polypeptides, 3714, 16742, 23546 and 13887 and nucleic acid molecules encoding them. The protein kinase polypeptides can be expressed by standard recombinant methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides are useful for diagnostic and screening methods to identify subjects (at risk of) having cancer or cellular proliferation and/or differentiation disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and modulators are useful for the treatment of cancer, particularly colon cancer or cellular proliferation and/or differentiation disorders. Other disorders associated with 3714, 16742, 23546 or 13887 expression or activity that can be treated include bone related disorders, inflammatory disorders, autoimmune diseases, cardiovascular disorders and liver diseases. The present sequence represents a human protein kinase polypeptide 3714.

XX Sequence 783 AA;  
 SQ  
 Query Match 20.2%; Score 831.5; DB 22; Length 783;  
 Best Local Similarity 30.3%; Pred. No. 1.2e-47;  
 Matches 249; Conservative 99; Mismatches 281; Indels 193; Gaps 26;

QY 31 VGPYRLKTLGKQGTGLVKLVGHVCHITGQKVAIKIVNEKLSSEYIAMKVERITALKLEH 90  
 DB 24 VGFYDIERTLGKGNFVAVKLARHRTVTKYAIKIDTRIDSSLEKTYREYQMLKLNH 83  
 QY 91 PNYLKLHDYENKKYLYLVLEHVSSELFDYLVKKGRLTPREARKFRQIYVSLDFCHSY 150  
 DB 84 PHIKLYQWETMDMYIYVTEFAKNGEMFDYLSNCHLSSEARKKFWQILSAVEYCHD 143  
 QY 151 SICHRDLKPEINLLDEKNNIRIDFGNASIQVDSLSLETGCGSHYACPEYIKKEKIDGR 210  
 DB 144 HLYHRDLKTEINLLDGDMDIKLADPFGFNFKSGEPLSTWCGSPVYAPAEVFEKEKEGP 203  
 QY 211 RADMWSGVILFALLVGALEPDDNLRQLLEKYRGVFNHPIPPCQSLIRGIMEVER 270  
 DB 204 QLDIWSLGVVLYLVYCGSLPFDGPNLPTLRQVLEGRFRIFPMSQCESTIRMLVYDP 263  
 QY 271 EKRLSEIQKHPWYLGKHEPDCLEPAG-RVAMRSLPSN-GEIDPVLESMAISGC 328  
 DB 264 ARRTIAQIRHRRM-----RAEPCL-PGPACPAFSAHSYTNIGDYDEQALGIMQTIGV 317  
 QY 329 FRDRELRHRESENEKMIYLLDRKERYPSCEDODLPKNDVDPKRRKRVDSMLS 388  
 DB 318 --DROKTVSLONSYNNRAIYLLERLKEYRNAD-----CA 354  
 QY 389 RRG-KRPERKSMELVISTDAGGSPVPTRRALEMAQHSRQSVGASATG--LSSS-- 443  
 DB 355 RGPAPQPPRSSDLSGLEVPQEGSLTDPFRPALCQPTQIVSQVLAEMDELQSLQ 414  
 QY 444 -----PLSSPRSPVSESP-----EPGAGDARGGGSTKTQTLPSKPPRGCGAQEQ 491  
 DB 415 WPLFPVPVDSGCGVFRPRPVSPSSLLDTALISEARQGPGLSEEDQDES--PSTGRHRT 473  
 QY 492 PPPSARSTPLGP-----PGSPRSSGGTPLL-----518  
 DB 474 LAEVSRLSPPLARPCIVBSPSTTASPAEGSSDSCLTFSSKSPAGLSGTPAIOGLIGAC 533  
 QY 519 SPLHPPRASP-TGTPEGTPPPSPGCVGA-----AMSRRLN 554  
 DB 534 SPVRL--ASPFILGOSQATVLAQGLGAVLLPVSEFGRRASDTSLTGLKAFRQQLR 591  
 QY 555 SIRNS--FLGSPRRH---RRKMOVPTAEEMSS-----LTPESPPEL-----AKRSMFGN 598  
 DB 592 KTRTRKGFGLNKKIGLARQVCPVPSRASRGSLSPFHAPAQSPGLHGAAGSRGEM--- 648  
 QY 599 FLSLDEEQIFVLKDKPLSIKADIVHAFSLPSLSHVSLSQTS- FRAEYKASGGPSV 657  
 DB 649 -----SLLEEVLEQQRLLQLQHHPAAAP-----671  
 QY 658 QKPVRFQVDISSEGEPEPRR-----DSGGGGIYVYTF---LISGPRRRFRKRYVE 707  
 DB 672 -----GCSQAPQAPAPAPVYIAPCDGPAAPLPSTLLTGLPLPPLQLTGASPV 721  
 QY 708 TIOAQLSTHDPQSVQALDEKNGAQTTPAGAPPSRLQPPG 749  
 DB 722 ASAAQLDTH-----LHIGTGPALPAVPPPRRLARLAPG 755

RESULT 12  
 ID AAB85786  
 AA AAB85786 standard; Protein; 783 AA.  
 AC AAB85786;  
 DT 29-OCT-2001 (first entry)  
 DE Human kinase PKIN-5.

XX PKIN: kinase; cytosolic; immunosuppressive; immunostimulant; human;  
 KW antileukosclerotic; cardiact; gene therapy; antisenescence therapy.  
 OS Homo sapiens.  
 XX MO200160991-A2.  
 XX 23-AUG-2001.  
 XX 16-FEB-2001; 2001WO-US05240.  
 XX 17-FEB-2000; 2000US-0183682.  
 XX 02-MAR-2000; 2000US-0186559.  
 XX 09-MAR-2000; 2000US-0186506.  
 XX 17-MAR-2000; 2000US-0189998.  
 XX 30-MAR-2000; 2000US-0193851.  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX Tang YF, Buford N, Gandhi AR, Patterson C, Khan FA, Yue H;  
 PI Hafila A, Shih LL, Tribouley CM, Yao MG, Burrill JD, Marcus GA;  
 PI Zingler KA, Lu DM, Bandman O, Pollock JL, Griffin JA, Thornton M;  
 PI Nguyen DB, Lal P, Walsh RT.  
 XX MPI: 2001-514771/56.  
 XX N-PSDB; AAH6213.  
 XX Isolated human kinase polypeptides useful in the diagnosis, treatment  
 PT and prevention of cancer, immune disorders and disorders affecting  
 PT growth and development.  
 XX Claim 1; Page 107-109; 126pp; English.  
 XX The invention provides human kinases (PKIN) and polynucleotides encoding  
 CC PKIN. The PKIN polypeptides can be expressed using standard recombinant  
 CC methodology. The PKIN polypeptides, polynucleotides, modulators and  
 CC specific antibodies are useful in the diagnosis, treatment and prevention  
 CC of cancer, immune disorders, disorders affecting growth and development,  
 CC atherosclerosis, and other cardiovascular diseases, and lipid disorders  
 CC and in the assessment of the effects of exogenous compounds on the  
 CC expression of nucleic acid sequences of human kinases. The present  
 CC sequence represents a human PKIN-5 polypeptide.  
 XX Sequence 783 AA:  
 SQ  
 Query Match 20.2%; Score 831.5; DB 22; Length 783;  
 Best Local Similarity 30.3%; Pred. No. 1.2e-47;  
 Matches 249; Conservative 99; Mismatches 281; Indels 193; Gaps 26;  
 QY 31 VGYRIEKTIGKGTGIVKGCITGOKVAIKTVNREKLSVLMKVERETAILIKLIEH 90  
 DB 24 VGYDIERTIGKGNFAVVLARHRVTKVOAIIKIDTRLDSSMLEKIYREVQIMKLINH 83  
 QY 91 PHTKLHDVYENKKYLYLVLEHVSGLFYLKKGRLTPREARKFFQIYSAIDFCHSY 150  
 DB 84 PHTIKLYOWMETDMKLYITIEPAKNGEMFDLTSNGHISENARKKQWLLISAEYCHDH 143  
 QY 151 SICHRLLKPEMLLDKKNRIRIADFGMASLQVDSLETSCGSPHYACEYIKGKPYGR 210  
 DB 144 HAYHRLTKTENLLDGMOKLADFGFNFKYKSEPLSTWCGSPRYAPAEYEGEGYGP 203  
 QY 211 RADMSCGVILFLLVGLALPFDDNRLQLEKVKRGVPHMHPITPDQCSILRGNIIEVP 270  
 DB 204 QLDIWSIGVLLVYVCGSLPFDGPNLPTLRORVLEGRFRIFPFMSOCESLIRMLVYDP 263  
 QY 271 EKRLSEIOIKHMYLGGKHEPPDCEPAPG-RVAMRSLSN-GEIDPDVLESASIGC 328  
 DB 264 ARITITAIQIHOHMM-----RAEPL-PGPACPAFSASHSTISNLGDDDEQALGIMQTLGV 317  
 QY 329 FDRERLARELSEENOEKMIYLLDRKERYPSCEDODLPPRNDVPPRRKRVDSPLMS 388  
 DB 318 --DRQRTVESLQSSSYNHFAIYLLERLKEYRNAQ-----CA 354

QY 389 RHG-KRRPERKSMELVISTDAGGGSPVTRALEMAHQSRSRVSGASTG--LSSS-- 443  
 DB 355 RPPANQPRRRSDLSGLFVQEBLSTDFRFPALLCQPPQIVSGVLAENDCELOSSLQ 414  
 QY 444 -----PLSFRSPVFSFSP-----EPGAGDEARGGSGPTSKTQTLPSRPGGAGEQ 491  
 DB 415 WPLFFPYDAGSGVFRFRPVSPSSLDLTALISEARQGGLEEDDTQESL-PSSTGRHRT 473  
 QY 492 PPPPSARSTPLPGP-----GSPRSSGCTLH----- 518  
 DB 474 LAEVSTRSLPFLTAPCIYVSPSTTASPAEGTSSDSCLTFPSASKSPAGLSGTATOGGLGAC 533  
 QY 519 SPLHTPRASP-TGTPGTTPPPSPGGVGA-----AMRSRLN 554  
 DB 534 SPVRL-ASPFISQSASTPVLAQGGIGGANLVVSPQEGRRASDTSLTQGLAKFRODLR 591  
 QY 555 SIKNS--FLGSPRPH--RRKQVPTAEISS--LTPSSPEL-----AKRSWFGN 598  
 DB 592 KTRIRKGLINKIKGLARVQCPASRASRGSLSPHAPASPGHAGAGSGREGW--- 648  
 QY 599 FISLDEEQIFVLKOKPLSLKADIYHAFILSTLSHVSLSQTS-FAEKKASGSPVF 657  
 DB 649 -----SLSEVLEQQRLLQLQHHPAAP--- 671  
 QY 658 QKEVRFQVDSISSEGEPEPSRR-----DSGGGGGIYSVFT-----LISGSRREKRYVE 707  
 DB 672 -----GCSQAPQAPAPAFVIAPCDGCAAPLSTLISGLPILPPLLTQTAGSPV 721  
 QY 708 TIOAQILSTHDQPSVQALADEKNGAOTRPAAPRSLQPPG 749  
 DB 722 ASNAQLDTH-----LHIGTPTALPAVPPRLRLAPG 755  
 RESULT 13  
 AAM78349  
 ID AAM78349 standard; Protein; 768 AA.  
 XX AAM78349;  
 AC 06-NOV-2001 (first entry)  
 DT Human protein SEQ ID NO 1011.  
 DE Human protein SEQ ID NO 1011.  
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 OS Homo sapiens.  
 XX MO200157190-A2.  
 XX 09-AUG-2001.  
 PF 05-FEB-2001; 2001WO-US04098.  
 XX 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YF, Liu C, Drmanac RF, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX MPI: 2001-476283/51.

DR N-PSDB; AAK51482.  
 XX Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 XX  
 PS Claim 20; Page 3241-3243; 6221pp; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAK51456-AAK53435) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAK60020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 CC  
 XX  
 XX

SO Sequence 768 AA;

Query Match 20.2%; Score 828.5; DB 22; Length 768;

Best Local Similarity 31.1%; Pred. No. 1.9e-47;

Matches 247; Conservative 110; Mismatches 260; Indels 177; Gaps 26;

QY 30 YVGPYRLKTLGKGGQGLVGLVGVCTGOKVAITVNRKLSVSLMKVEREIALKLE 89  
 DB 40 HIGNYRLKLTIGKGNFPAKVLARHILTGKEVAVKIIDKTQUNSSLOKLFREVRIMKVLN 99  
 QY 90 HPHVLAHDYENKKYLYLVLEHVSGLGELFDYLVKGRLLPTEKARKFFQVIALDFCHS 149  
 DB 100 HPMIVKLFEEYETFEKTLVLMVEYASGEVFDYLAHGMKKEAKRQVIAVAYOYCHQ 159  
 QY 150 YSICHRDLKPEENLLDKENNRIRADFGMASIQVDSLETSCGSHYACPEYIKKEKIDG 209  
 DB 160 KFTVHRDLKPEENLLDKENNRIRADFGMASIQVDSLETSCGSHYACPEYIKKEKIDG 219  
 QY 210 RRADMMSGYILFALVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLV 269  
 DB 220 PVDVMSGLYILVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLV 279  
 QY 270 PEKRLSLEQIQKHPWLGKHEPD---PCLPAPGRVYAMSLPENGDELVDYLSMASL 326  
 DB 280 PSKRGTEQIMKDRMMVNG--HEDELKRYVEPLPDYK-----DPRTELAMVM 326  
 QY 327 GCRDRRLRLRELSEENOEKMYIYLLDKKE----- 359  
 DB 327 G--YTRREIDSLVGORYN--EVMATYLLGLYKSELEGGDTITLKPRPSADLTNNSAHPH 383  
 QY 360 -RXPSC-----EEOADLP-----RNDVDPKRVDS----- 384  
 DB 384 TRRVSANANKORFSDQAGPPIPTNSYSKKTOSNNAKENKRPEDRESGRASSTAKYV 443  
 QY 385 ---PMLSRGKRRRPERKSMSEVLSITDAGGSGSPVTRALEMA---QHSQRSVSG--A 436  
 DB 444 ASPLPGLERK-KTPTPTSTNSVLSITSTSRNSRNSPLERASIGQASIQNGKDSLWPGSRA 502  
 QY 437 STGLSSPLSPR-----SPVFSFPEPGAGDEKRGSGSPSTKQTLPKSRP----- 483  
 DB 503 STASASAAVSAARROHOKSMSASVHPKASGLPPTESNCEVPRSTAPORVVASPSAH 562  
 QY 484 ---RGAGAGAO---PPPSARSTPLPGRPPSGSGGTPPLSLHTPTASPTGTGTPP 537  
 DB 563 NISSSGAPPRTPNPRGVSSRSTFHAGOLROYROQNLPT-----GVTP 606  
 QY 538 PSBGVGGAAMSRSLNRSIRSLGSPFRHRRKMQVPTAE--MSLTP----- 584  
 DB 607 ASFGSHQGR--RGASGSIFSFK--TSKFRVARNLNEPESKDRVETLRRHVVGSGGDKK 662  
 QY 585 ---ESSPELAKNSW--FGNTISLDKEQIYLVLDKPLSLIKADIYVHAFSLPSLSHSLV 639

DB 663 EEFREAKPRSLRFTWSKRTTSSMEPMEMREIRKVLNDANSQSELHKKYML--CMGTTP 720  
 QY 640 SQTSFAEYKASGSPSVFQKPVRFQVNDISSSEGEPSRDRSGGGGGLYSTFTLLISGS 699  
 DB 721 GHEDF-----VQMEVECK-----LPRSLNSG-----VRFRIISGTS 752  
 QY 700 RREKRVETIOAL 713  
 DB 753 MARKNIASKIANEL 766

RESULT 14  
 AAE19050

ID AAE19050 standard; Protein; 691 AA.

AC AAE19050;

DT 18-JUN-2002 (first entry)

DE Human PAR-1B alpha protein.

KW Human; Dishevelled associated kinase; PAR-1 kinase; enzyme; PAR-1A;

KW PAR-1B alpha; PAR-1B beta; PAR-1C; cancer; hyperproliferative disease;

KW antisense therapy; Dsh.

OS Homo sapiens.

PN WO200210402-A2.

PD 07-FEB-2002.

PF 30-JUL-2001; 2001WO-US233981.

PR 28-JUL-2000; 2000US-221860P.

PA (CHIR) CHIRON CORP.

PI Sun T, Feng J, Reinhard C, Pantl WJ, Williams LV;

DR WPI: 2002-206192/26.

DR N-PSDB; AAD30397.

PT Isolated nucleic acids encoding the human and Drosophila Dishevelled

PT associated kinase, referred to as PAR-1 kinase, useful for identifying

PT modulators which can be used for modulating, inhibiting or preventing

PT the growth of cancer cells -

PS Claim 7; Page 283-284; 287pp; English.

XX The invention relates to Dishevelled (Dsh) associated kinase, referred

XX to as PAR-1 kinase in Drosophila and human homologues of PAR-1 referred

XX to as PAR-1A, PAR-1B (alpha and beta) and PAR-1C. The invention also

XX provides nucleic acid molecules encoding such proteins. PAR-1 activates

XX the Wnt pathway and is required for Wnt signalling in mammalian cells.

XX The PAR-1 modulators can be used to reduce the expression and/or

XX biological activity of PAR-1. They are useful for modulating, inhibiting

XX or preventing the growth of cancer cells. They are also useful for

XX treating other diseases of hyperproliferation. Sequences of the invention

XX are used in antisense therapy. The present sequence is human PAR-1B alpha

XX protein.

SO Sequence 691 AA;

Query Match 20.2%; Score 828; DB 23; Length 691;

Best Local Similarity 30.9%; Pred. No. 1.8e-47;

Matches 232; Conservative 115; Mismatches 261; Indels 142; Gaps 24;

QY 30 YVGPYRLKTLGKGGQGLVGLVGVCTGOKVAITVNRKLSVSLMKVEREIALKLE 89  
 DB 16 HIGNYRLKLTIGKGNFPAKVLARHILTGKEVAVKIIDKTQUNSSLOKLFREVRIMKVLN 75  
 QY 90 HPHVLAHDYENKKYLYLVLEHVSGLGELFDYLVKGRLLPTEKARKFFQVIALDFCHS 149

```

DB 76 HPIVILFEVIEFEKTLVMEVYASGEVFYLAHGMKEKEARAKROIYSAVOYCHQ 135
QY 150 YSICHRLDKPENLLDKENRIADFGMASLQVDSLLETSCGSPHYACPEVYKGEKYDG 209
DB 136 KFIYHRLKAEENLLDADMNIKIAIDFGFSNEFTGKNKLDTCGSPHYAPALFQCKKYDG 195
QY 210 RRDWMSCGVILFALVAGALPDEDDNLRQLEKRYGFHHPFPPCQSLRGMIYEY 269
DB 196 PEVDWMSLVGILVTLVSGSLPFDQNLKELREYLRGKYRIPFYSTOCENLKLKFLILN 255
QY 270 PEKRLSLEJOIKHPWYVYGGKHEPD---PCLPAPGRVYAMSLPENGELDPVLSMASL 326
DB 256 PSKRGTLEJOIKMKRMVYV---HEDELKPYVPLDYK-----DPRTELAMSM 302
QY 327 GCFRDRERLRELRSEBNOEKMTYLLDRKERYPSCEDDLPVRNDVDPKRVDP- 385
DB 303 G---YTRIEDIDSLVQRYN---EVMATYLLGYKSELEGDITLKRPRPADLNNSSAPSPS 359
QY 386 -MISRHKRRPERKSMVSLTIDAGGGSPVTRALE---MAQHSQSRVSGASTGLIS 441
DB 360 HKVQRSVSNPKORF-----SDQAGPALPTNSYSKKTQSNNAENKRPEDRESGRK 412
QY 442 SSPSPSPVSPVSPSPPGAGDEARGGSGPTKTQTLPSRGRGAGGEQPPPSARSP 501
DB 413 AS---STAKVPA---SPLPGL---ERKKTPTPTNSVLTSTN-----RSRNSP 453
QY 502 L--PGPPSPRSSGTPDLSPATPRASPT-----GTPGTT----- 535
DB 454 LLEBASLGQASIQNGKSDTAPORVVAPSPSAHNISSGADRTNFPGRVSSRSTFHAGQ 513
QY 536 -----PPSPGGVGGAAMRSRLNSTRNSFLSPRRHRKMYPTAEF--MS 580
DB 514 LRVQDQNLPGVTPASPSGHSGR--RGASGISFSEK--TSKVRNRLNMPESKDRE 569
QY 581 SLTP-----ESSPELAKRSW-FGNFISLDEKEQFLVKDKPLSIRAD 623
DB 570 TLPHVYSGGNDKEKEFPRKRLFTSMKTTSMENEMRELIRKYVDANSQSE 629
QY 624 IVNAFSLPBLSHSVLSQTSFRAEYKASGGSPVOKPYRFOVDISSSGPEPSPRDGS 683
DB 630 LHKRYML--CMGTGPHEDF-----VQMEYVCK-----LPLRLSLNG 665
QY 684 GGGIVSTFTLISGSRFRKRYVETIAQL 713
DB 666 -----VRFKRISGTSMAFKNIAKIANEL 689

RESULT 15
AAE19051
ID AAE19051 standard; protein; 724 AA.
XX
AC AAE19051;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human PAR-1B beta protein.
XX
KW Human; Dishevelled associated kinase; PAR-1 kinase; enzyme; PAR-1A;
KW PAR-1B alpha; PAR-1B beta; PAR-1C; cancer; hyperproliferative disease;
KW antisense therapy; Dsh.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Misc-difference 274
FT /note- "Encoded by AGG of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 275
FT /note- "Encoded by TGC of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 276
FT /note- "Encoded by CTG of the inverse complementary

```

```

FT Misc-difference 277
FT /note- "Encoded by ACA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 278
FT /note- "Encoded by CTT of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 279
FT /note- "Encoded by TTG of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 280
FT /note- "Encoded by GAC of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 281
FT /note- "Encoded by GAA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 282
FT /note- "Encoded by TTC of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 283
FT /note- "Encoded by TTT of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 284
FT /note- "Encoded by AAA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 285
FT /note- "Encoded by GAG of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 286
FT /note- "Encoded by TAA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 287
FT /note- "Encoded by GAA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 288
FT /note- "Encoded by TTA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 289
FT /note- "Encoded by GGG of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 291
FT /note- "Encoded by TTC of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 292
FT /note- "Encoded by TCT of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 293
FT /note- "Encoded by CCG of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 294
FT /note- "Encoded by TGA of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT
FT Misc-difference 295
FT /note- "Encoded by AAT of the inverse complementary
FT strand of the sequence shown as SEQ ID NO:8 (AAD33068)."
FT

WO200210402-A2.
XX
XX 07-FEB-2002.
XX
XX 30-JUL-2001; 2001MO-US23981.
XX
XX 28-JUL-2000; 2000US-221860P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Sun T, Feng J, Reinhard C, Fantl WJ, Williams LT;
XX WPI; 2002-206192/26.
XX N-PSDB; AAD30398, AAD33068.
XX Isolated nucleic acids encoding the human and Drosophila Dishevelled
XX associated kinase, referred to as PAR-1 kinase, useful for identifying

```





QY 440 LSSPLSPSPV-----FSFSPGAGDEARGGSPSTKT-----QTLPS 480  
 DB 411 ---SETPESRSSTRDLFGSSSSGYSARAGEDNDGRSASRSTSYHYYPQVADPQTLAE 467  
 QY 481 -----KCPGCGAGE-----QPPPSANSTPLPGPGSPRS 511  
 DB 468 AARHVDAQERRRSRSGRSSRKRKESKRSKSDKASASSSSCKNDASTSTSPKRYSPSYMS 527  
 QY 512 SGGTPLHSPDLPRASPTGPTPTPPSPGCGVGAAMRSLNIRNSFLGSPFHRKK 571  
 DB 528 ESIVVSSSTMNSTNSSTNSLIGNSQTSTIGTSG--PWRSKLNNIKSFLGTPFHRKKM 585  
 QY 572 QVPTAEEMSLTPE--SSPELAKRSWFGN--ISLDKEOIFLYLKDKPLSSIKADIYHA 627  
 DB 586 SNGTASDESDSQMIDTDTLVKRSKWFGLASSMSVERDTHCVQGTKLNSIKAEILTRA 645  
 QY 628 FLSTPLSHSVLSQTSFRAEYK--ASGSPSVFQKPVAFQVDISSEGPSPRRDG--SGG 684  
 DB 646 FLQIHELHSHSVGQNEFRYKRGKPTVGGSVFSKIGKMDIIT-----PSPQOVVIAE 699  
 QY 685 GGIYSVTFLLISGSPRRFRKRVETIOALSTFDQSPVOALADEKNGA--QTRP 736  
 DB 700 TPTVYVQFVLIAAGPVARRFKLVHEHLSIL-----QNSTQGRADRGQAALMVRP 748

## RESULT 2

S37928  
 probable purine nucleotide-binding protein YKL101w - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YK1453  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 24-Sep-1999  
 C:Accession: S37928; S39084  
 R:Cheret, G.; Fukuhara, H.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Pallier, C.; Puzo  
 submitted to the Protein Sequence Database, March 1994  
 A:Reference number: S37920  
 A:Accession: S37928  
 A:Molecule type: DNA  
 A:Residues: 1-1518 <CHR>  
 A:Cross-references: EMBL:X71133; NID:9486168; PID:CAA81941.1; PID:9486169; MIPS:YKL101w  
 A:Experimental source: strain S288C  
 R:Pallier, C.; Valens, M.; Puzos, V.; Fukuhara, H.; Cheret, G.; Sor, F.; Bolotin-Fukuhara  
 Yeast 9, 1149-1155, 1993  
 A:Title: DNA sequence analysis of a 17 kb fragment of yeast chromosome XI physically loc  
 protein kinases.  
 A:Reference number: S39084; MUID:94078677; PMID:8256524  
 A:Accession: S39084  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1518 <PAL>  
 A:Cross-references: EMBL:X71133; NID:9431205; PID:CAA50456.1; PID:9431215  
 A:Experimental source: strain S288C  
 C:Genetics:  
 A:Gene: SGD:HS11  
 A:Cross-references: SGD:S0001584; MIPS:YKL101w  
 A:Map position: 11L  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
 C:Keywords: ATP; P-loop; purine nucleotide binding; serine/threonine-specific protein ki  
 F:79-369/Domain: protein kinase homolog <KIN>  
 F:79-86/Region: nucleotide-binding motif A (P-loop)  
 F:87-95/Region: protein kinase ATP-binding motif  
 F:85/Binding site: ATP/GTP (Lys) #status predicted

Query Match 20.4%; Score 836; DB 2; Length 1518;  
 Best Local Similarity 28.5%; Pred. No. 1e-22;

Matches 256; Conservative 135; Mismatches 261; Indels 246; Gaps 29;

QY 31 VGPYRLKLTGKQGTGLVKGHCITQQAIVKIVNREKL----- 70  
 DB 78 VGPWKGTGKTLGKSSGRLAKNMETQLAIAIKIVPKKAVHCSNNGTVPNSYSSSMVT 137  
 QY 71 ----SSSVLMK-----VERIATILKLEHPPHVLKLVHYENKKTLYLVLEHVS 114  
 DB 138 SNVSPSIASRHSNHSQTNPDYGIERTIVIMKLISHVNMALFEWENKSELYLVLETVD 197

QY 115 GGEFLDYLVKRGRLTPEAKRFEQIVSALDFCHSYSLCHRLDKPENLLDEKN--NRIA 173  
 DB 198 GGEFLDYLVKRGRLTPEAKRFEQIVSALDFCHSYSLCHRLDKPENLLDEKN--NRIA 257  
 QY 174 DFGMASLOYDLSLETSCGSPHYACEVIKGEYKDYGRADMGCGVILFALLVGALEPDD 233  
 DB 258 DFGMALDLPENLLKTSKSPHYASPEIYMGPRHGGSPSYMGCVILFALLVGHLPFND 317  
 QY 234 DNLRLLEKVGKGVPMHPIPPDQSLRGMEVEGERKSLDEQLOKHW----- 284  
 DB 318 DNLRLLEKVGKGVPMHPIPPDQSLRGMEVEGERKSLDEQLOKHW----- 377  
 QY 285 -----YLGRHEPDCLPEAPGRVAMRSLPSNGELDDVLESNASLGFEDRE 333  
 DB 378 NKVLKMRKMDNARKSNSD--LHLNNVSPSIVTLHSGKEIDSLRSQILMHGVSRE 435  
 QY 334 RLHRELSEENDEKMTIYLLDRKERY-----PSCEDDLPFRNDVPPRKVDSPMS 388  
 DB 436 LITAKLQPMSEKLFYLLQYKQRHSISLSSSEKKSATSSVNERIEVASTAN 495  
 QY 389 RGRKRPERKSMELVSLTDAGGGSPVPTPRALMAQHSO-----RSRVSAGSTGL-- 440  
 DB 496 NTGLR-----SENNDYKTLHSLFT--HSDDTGVNOMNAITGVNTEINA 537  
 QY 441 -----SSSPLSPRS-----PVFSFSPGAGDEARGGSPSTKTQT 477  
 DB 538 PVLAQKSPSINTLSQPESDKAEAVTLPPALPIFNAS-----SSRIFRNYSITSS 590  
 QY 478 LPSKRGAGGAGEQPPPARST-----PLPSPGSP-----RSAGTTP--LHP 520  
 DB 591 RSRSLRLNSRLSLSASYSRELVHONEMLPOLPSPSPSYSLRRRIHASPSTKTHKS 650  
 QY 521 LHTPRASPTGPTPTPPSPGCGVGAAMRSLNSI--RNSFLGSPRRRRKMQV----- 573  
 DB 651 LSRKNIAAT-----VAARTLQNSAKRSLYSLSQTSKININDLVF 694  
 QY 574 -----PTAEEMSLTPESSPE-----LAKRSWFGNFI--LDKEOIFLYLKDK-- 615  
 DB 695 DDLPSKPKASEVNVKSEPHSLSDSDDFELICQIIFGNALDRILEEED-----NKKER 749  
 QY 616 -----PLSTIADIVHAFSLPSLSHVSLSQTSFAEYKASGSPS--- 655  
 DB 750 DTRQRONDTKSSADTFTTSVSTKNENEGEYP---TYIEKNQFMSTKPSSEMGSL 805  
 QY 656 ---VFQKPVAFQVDISSEGPSPRRDGSGGGIYSVFTLLSGPARRKRVETIQ-A 711  
 DB 806 SPPIFEK-----ENTLSSYLEQKPKR-----AALSDITNSFKNMKNQDEMRITKQORE 856  
 QY 712 QLTSTHDQPSVOALADEKNGAQTRPAGAPPSLQF-----PGRPPDELSSPPR 761  
 DB 857 QLOKKNDRPS-----PLKPIQHDELRYVNSLPLNDQKPSLSLDPRR 896

## RESULT 3

G01025  
 serine/threonine protein kinase - human

C:Species: Homo sapiens (man)  
 C>Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 16-Dec-1998  
 C:Accession: G01025  
 R:Nayyar, E.  
 submitted to the EMBL Data Library, April 1996  
 A:Reference number: H00564  
 A:Accession: G01025  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-745 <NAV>  
 A:Cross-references: EMBL:X97630; NID:q1310674  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom  
 F:18-271/Domain: protein kinase homolog <KIN>

Query Match 20.2%; Score 828; DB 2; Length 745;  
 Best Local Similarity 31.0%; Pred. No. 1.1e-22;



Db 621 SLAPPSTKSSNTTQIAIPLFNNRSTATSSAOPSTGTRKRIADKGRIPLNSTA 680  
 QY 544 VGGAAWRSLNIRNLSFLGSPFRHRKM-----QVPAEEMSL--TP----- 584  
 Db 681 VGHRTATGAVANNNGIPSHDHQAQQOQYMNOLSTSTMSKLINTKPAAGTAISSSS 740  
 QY 585 -----ESSPELAKRSMFNGF----- 599  
 Db 741 SSSATSTAPLAKSSGQISIAPTPEVYREDDENNENQNNVPLIGVGQTPSPAVQPT 800  
 QY 600 ---ISLDEQIFLVLKDKPLSSIRADIYHAFILPSLSHVSLOTSPRAEYKASGSPS 655  
 Db 801 EDATSSSDKEQOQKASSETPKES-KPSMHQSPSMPPSQMTAMELSKLSSEGQGTGPT 859  
 QY 656 VYQKVRVQVYDSSSGSPSPRRDSSGGGIGYSVFLLISGPRRRKRVETIQALLS 715  
 Db 860 V-----ATGPG--PQR-----ATSSQMSRSATTTISAMNGA 887  
 QY 716 THDQPSVALADEKNGAQTRPAGAPPRSLDPPGRDPELSSPRRGPDPKDL 769  
 Db 888 SSGGAAAAAATNOLSGASPSTGASSGQYHPKA-----PSSSSSTNPMPHQHL 937

## RESULT 5

S27966  
 probable serine/threonine-specific protein kinase (EC 2.7.1.-) - human  
 N.Alternate names: protein p78  
 C.Species: Homo sapiens (man)  
 C.Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 04-Mar-2000  
 C.Accession: S27966  
 R.Maheshwari, K.K.; Som, S.; Parsa, I.  
 submitted to the EMBL Data Library, January 1992  
 A.Description: Sequence of a cDNA encoding 78kD marker protein lost in chemically induced  
 F:54-307/Domain: protein kinase homology <RIN>  
 F:62-70/Region: protein kinase ATP-binding motif  
 A.Accession: S27966  
 A.Molecule type: mRNA  
 A.Residues: 1-713 <MAN>  
 A.Cross-references: EMBL:M80359; NID:g189511; PIDN:AA59991.1; PID:g189512  
 C.Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
 C.Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
 P:54-307/Domain: protein kinase homology <RIN>  
 F:62-70/Region: protein kinase ATP-binding motif

Query Match 19.4%; Score 797; DB 2; Length 713;  
 Best Local Similarity 32.0%; Pred. No. 1.3e-21;  
 Matches 237; Conservative 112; Mismatches 254; Indels 138; Gaps 28;

QY 30 YGPPYLEKTLGGQGLVNLGYHCTGQVAKIVNRKISSEVLMKVEEILITKIE 89  
 Db 52 HIGNYRLKTKIGKGNPAKIVLAHRIILTGREVALKITDKTQLNPLSLOKLFREVRIKTLN 111  
 QY 90 HPHVTLKHVYENKKYLYLVLEHVSSEGLDYLVKKGRLLPKARKKFFQVSLALDFCHS 149  
 Db 112 HPIVTLFVETIETQKLYLIMEYASGGKVFYVAHGRMKERKSRQIVSAVOYQHO 171  
 QY 150 YSICRDLKRENTLLDEKNNIRIADFGMASLQVDSLLETSCGSPHYACPEYIKGEKIDG 209  
 Db 172 KRIYHDLKAEENLLDADNMIKIADFEFSNEFTVGKIDLPFCGSPHYAAPLFOCKKIDG 231  
 QY 210 RRAMMSCGVITALLVGAALPDDDNLRQLEKVRGVFHHHPFPCCOSGLRMIMEVE 269  
 Db 232 PEVDVMSLVGLVTLVSGSLPPDQGNLKELEKRYLRKRIPTFWSIDCENLAKFVLIN 291  
 QY 270 PEKRLSLQEQKHPWYLGKHHPD--PCLEPAGRRVAMSLSPSNGEL--DDPVLSM 323  
 Db 292 PIRKGLTEQIMKRWIAG--HEDEELKPYVRP-----ELDISDQKRIDIM 335  
 QY 324 ASIGCFRDLRRLHRELSEENQ--KMTY-----YLLDKREYVSCEDQLPPRND 374  
 Db 336 VGMG-----YSOEIQSLSKMKKDEITATYLLGRR-----SSEVRPSSD 376  
 QY 375 VD-----PPKRVDSPLSRKGRRRPERKSMELVSLTDAGGGSP-----VPRRALEMAQ 425

Db 377 LNSTGQSPHHKVRQSVSSQKOR-----YSDHAGQIPSVAVYPRKNSQSTAD 426  
 QY 426 HSQR-----SRVSGASTGLSSPLSSPRPVFSFSPGAGDEAAGGSPSTQTLPS 480  
 Db 427 SDLKEDGISRSKSTSAVGGK---IAPASPMGNASNNKADIP-----RKSSIVS 478  
 QY 481 RQPRGAGGEQPPPSANSTPLPGPPGSPRSSGCPPLHSPLHTPRASPTGPTPPSP 540  
 Db 479 SNTASGM-----TRNTYV-----CSERTTDDR--HSVIQNGKENST-IPDQRPVAS 524  
 QY 541 GGVGGAAMRSL-----NSIRNLSFGSPFRHRKQVPTAEMSLTDESSPELAKRS- 594  
 Db 525 THSISAAATPDRIRPRTGASSTFHQPR-EKRTATYNGPPASPSSHPAPLSQTRSR 563  
 QY 595 ---WFGNFI-LDKEQIFLVLDK-----KPLS-----SIK-----ADIVHAFSLIP 632  
 Db 584 GSTLFSKLTSLTIRSNVSAKQDENKRAKRSLLFTSMKRTSSMDPDMMRETRKVL 643  
 QY 633 SLHSVLQTSFRAEYKASGSPSVQKPYRFQYDVISSSGEPSPRRDSSGGGIGYSVF 692  
 Db 644 DANNCDYEQR-REFLLFCVHGDAENLVQWEMEVCK-----LPRLSLNG-----VRF 690  
 QY 693 TLISGSPRRKRVETIQAL 713  
 Db 691 KRISGTSIAFKNIASKANEL 711

## RESULT 6

JC1446  
 serine/threonine-specific protein kinase (EC 2.7.1.-) AK21 - Arabidopsis thaliana  
 N.Alternate names: protein Kinase SNF1 homolog  
 C.Species: Arabidopsis thaliana (mouse-ear cress)  
 C.Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 11-Jun-1999  
 C.Accession: JC1446; S58266; S66334  
 R.Ileguen, L.; Thomas, M.; Bianchi, M.; Halford, N.G.; Kreis, M.  
 Gene 120, 249-254, 1992

A.Article: Structure and expression of a gene from Arabidopsis thaliana encoding a prot  
 A.Reference number: JC1446; MUID:93013041; PMID:1339373  
 A.Accession: JC1446  
 A.Molecule type: DNA  
 A.Residues: 1-512 <LEG>  
 A.Cross-references: GB:M93023; NID:g166599; PIDN:AAA32736.1; PID:g166600  
 R.Themmler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.  
 submitted to the EMBL Data Library, May 1995  
 A.Description: Differential accumulation of the transcripts of 22 novel protein kinas  
 A.Reference number: S58266  
 A.Accession: S58266  
 A.Status: preliminary

A.Molecule type: DNA  
 A.Residues: 144-198 <THD>  
 A.Cross-references: EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PID:g928910  
 R.Themmler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.  
 Plant Mol. Biol. 29, 551-565, 1995  
 A.Article: Differential accumulation of the transcripts of 22 novel protein kinase gene  
 A.Reference number: S66334; MUID:96123233; PMID:8534852  
 A.Accession: S66334

A.Molecule type: DNA  
 A.Residues: 144-198 <TH2>  
 A.Cross-references: EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PID:g928910  
 C.Comment: This enzyme plays an important role in a signal transduction cascade regul  
 A.Genetics:  
 A.Gene: AK10; AK21  
 A.Introns: 64/1; 125/3; 186/3; 230/3; 292/3; 322/3; 350/3; 396/3; 475/3

C.Function:  
 A.Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threo  
 C.Superfamily: AMP-activated protein kinase; protein kinase homology  
 C.Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kin  
 F:17-271/Domain: protein kinase homology <RIN>  
 F:25-33/Region: protein kinase ATP-binding motif  
 F:48, 67, 142, 144/Active site: Lys, Glu, Asp, Lys #status predicted  
 F:147, 151/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 19.4%; Score 795.5; DB 1; Length 512;

Best Local Similarity 43.5%; Pred. No. 1.1e-21;  
Matches 157; Conservative 73; Mismatches 116; Indels 15; Gaps 5

OY	34	YRLKRTLGQGTGYLKVLCVHCITGOKAIVINREKLSEVL-MKYERETAILIKLEPH	92
Db	19	YKLTSTLGISGFVWKTAENHATLTHKVAIKILNRRKIKNMMEKEVKREIKIRLFMHPR	78
OY	93	VLKCHDYVENKKYLYLVLENHSGGEFLDYLKKGRLLPKARKFFROIVASALDFCHSYI	152
Db	79	IIRLXEVIETPTDIYLMVEYNSEGELDYLYEGRLOEDBARNFPOOITSGVRYCHRWNA	138
OY	153	CHRDLPENILLDEKNRIADIFGMSIAQGDSILLETSFGSPHYACPEYIKGEKYGRRA	212
Db	139	VHRDLKPENILLDSCKNVKLTADFGLSIMDHGHFKLTCSSPYTAAPEVYISGLTYAGPEY	198
OY	213	DMWCSGYTLRALVALGPDDDNLROLLEKVVAKRGVFMHPFIIPDCOSLRLGMIIEVEPER	272
Db	199	DWMSGVILVALCGTLPFPDENIPNLFFKIKIGIYTLPLSHLSPGARDLIPRLVYDPMK	258
OY	273	RLSLEJOXKHFWYGGKHNEPDLCPLEAPGKRVMRSLPSNGSELDPVLESMASTGCFFDR	332
Db	259	KVTTPETLRQHPFOA--HLRPYLAVPDP-----TVQQAARLIDEIIILEVINMGF--DR	308
OY	333	EHLRELRSSEENOEKMYIYLLDKRK-----YPSCEDDLP.RPNVDVDPKKRVDS.PMLS	388
Db	309	NHLIESLRNTQDNGFYTTYLLILDNMFRRASSGTLGAHFQETHMGTGTRPMHPAESVASPVSH	368
OY	389	R 389	
Db	369	R 369	

## RESULT 7

SNF1-related protein kinase [imported] - *Gulliardia theta* nucleomorph  
C:Species: nucleomorph *Gulliardia theta*  
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C:Accession: B90120  
R:Douglas, S.; Zanner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reiter,  
Nature 410, 1091-1096, 2001  
A:Title: The highly reduced genome of an enslaved algal nucleus.  
A:Reference number: A99082; PMID:11523671; PMID:11523671  
A:Accession: B90120  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-401 <DOU>  
A:Cross-references: GB:AF0803031; NID:g13794278; PIDN:AAK36655.1; GSPDB:GN00152  
C:Geneics:  
A:Gene: kin(snff2)  
A:Map position: 3  
A:Genome: nucleomorph  
Keywords: nucleomorph

Query Match	Score	DB 2;	Length
19.38;	793;		401;
Best Local Similarity	41.88;		
Best Local Similarity	50.00;		

	matches	137/	conservative	92/	mismatches	97/	indels	30/	gaps	8/
QY	29	QYGPYRLERLTKGGGGLVGLKGVHCITGQVAKIRVNRKLE-	-	SYLMVYREIRALIK	86					
Ddb	10	RFVGSYILDTLDTLVGAGKRYKLAKHSTKEVGVKIIKKDLPYDPSRLIKQIREISYWK			69					
QY	87	LIEHPVLKLDHYENKKYLVLVLEHVSGBELFDYLVKRGRLTPKEARKFPQIYALDF			146					
Ddb	70	LMHPHYIKIYDLEDSKYLFLITEVASKGLFYLVEKRLLENREALKFFHEIISGLEI			129					
QY	147	CHSYSTCHRLRKEENLLDEKNNIRIDFGMASLQVGDLSLETSCGSPHYACPEYIKGEK			206					
Ddb	130	CHNRHRCRHLKLENTLLDMKLIQIKIADFGLMASLISIPINIMLTKFCGSPHYASPEVYSNEP			189					
QY	207	YDGRADMDSCGYILFLLVGLALPFD-	-	DNIRQLLEKVRKGVHMFIPDQCSILRG	264					
QY	190	YNQIKADIMSCGIIILVALVYKGLPYDENNNRKLLENKIREPPRIPIPIHPDQCDLLRA			249					

[illegible]

OY	317	PDVIESMASTGCFEDRELRHLELNSEENOEMKIYYLLRLRKERYSPSCEDODLP	PRNDV	376
	:	:	:	:
Dd	300	PEITSFLPLQVLVDKDKTGLLGISTKKPSFLRVLYRQLERKRKMDSVRSNLI	--NEAN	356
OY	377	PPRRKRVDSFMLSRHG	392	
	:	:	:	:
Dd	357	EKKRK----	LFGKG	367

## RESULT

probable serine/threonine-specific protein kinase (EC 2.7.1.-) - cucumber  
N:Alternate names: SMF1-related protein kinase  
C:Species: Cucumis sativus (cucumber)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 20-Jun-2000  
C:Accession: T10449  
R:Gumpel, N.J.  
submitted to the EMBL Data Library, December 1996  
A:Reference number: Z17020  
A:Accession: T10449  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-504 <GUM>  
A:Cross-references: EMBL:Y10036  
A:Experimental source: cv. Masterpiece; cotyledon  
C:Function:  
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine-phosphate  
C:Superfamily: AMP-activated protein kinase; protein kinase homology  
C:Keywords: ATP, phosphotransferase; serine/threonine-specific protein kinase  
F:6-260/Domain: protein kinase homology <Kin>

Query Match	19.38	Score 792.5	DB 2	Length 504
Best Local Similarity	45.68	Pred. No. 1.4e-21		
Matches 151; Conservative	68	Mismatches 101	Indels 11	Gaps 4

QY 34 YRLKPLGKQDGYLKLGVHCITGQAKAIIVNREKISE-SVLKVERETALLKLEPH 92  
 Db 8 YKRGKLGISFQKVAEHALTGHHKAIKLNNRKIKLNDMEKQKREKIKILRLFMHPH 67  
 QY 93 VKLHDVYENKKYLLVLLEHVSGETLDYLVKKRGRLPKPAREFFROIVASALDFCHSISI 152  
 Db 68 IILRYEIEFTSPDIYVVMVEYKSGEFLDYIEVGRQLQEDBARNFPOQIIISGEVCHRNMY 127  
 QY 153 CHRDILKPEMLLDKKNRIRIADFGMSLQVGDSSILETSCGSPHIACPEYVTKGEYDGRRA 212  
 Db 128 VNHDLKPEMLLDKSKCVKRLADFGLSIMKMDGHFLTKTSCGSPYAAPEVYISGLIYAGPEV 187  
 QY 213 DMWSCGVILFALLVGLPEDDNLROLLEKVKRKGAVFMHPFIRPPDCQSLRGMIETVEPEK 272  
 Db 188 DMWSCGVILYALLCGFLPEDEDENIPMLFKIKKGITLPSHLSSGARRELIPSLVLVDDPK 247  
 QY 273 RLSLGEIOKHPWLTGSKHEDPCLEAPGRVAMRLPSNGELDPRVLESMASLGCFFRDR 332  
 Db 248 RILPIERLQHPMQA-HLRPIYAVRPD-----TMOAKKIDEDILQEVVMGF--DR 297  
 QY 333 EURLRELSEENOEMKITYLLIDRKERYPS 363  
 Db 298 NOLVESLNNRIQENATVATVYLLIDNFRVSS 328

## RESULT 9

```
protein H39E23.1 [imported] - Caenorhabditis elegans
C:/Species: Caenorhabditis elegans
C:/Date: 10-May-2001 #sequence_revision 10-May-2001
C:/Accession: G69287
#text_change 10-May-2001
R:anonymous. The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
```

A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology  
 A:Reference number: A75000; MIM:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/99C/C\_elegans/ and www.sanger.ac.uk/Projects/C\_elegans/  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: G89287  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1246 <SNO>  
 A:Cross-references: Gb:chr-V; PIDN:GAB09532.1; PID:93878100; GSPDB:GN00023; CESP:H39E23.  
 C:Genetics:  
 A:Gene: H39E23.1  
 A:Map position: 5

Query Match 19.3%; Score 791.5; DB 2; Length 1246;  
 Best Local Similarity 29.3%; Pred. No. 3.2e-21;  
 Matches 249; Conservative 126; Mismatches 305; Indels 169; Gaps 26;

```

2  SSGAKGCGGSP-AYHLPHPHPPOHAQ-----YVGPRLKLT 40
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7  SSGKRTVAVSPGDHYTHNRKSSSHARSTGCGMSRARRNDQVHGKYLKTI 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
41 GKGQGLVAVLVCITGOKVAIKIVNREKISEVLMKVERIALIKLIEHPVILKHDY 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 GKGNAKVKLAKHVIHGEVAIKIIDTALNPSSLOKLFREVKIMKOLDHPNIVKLYQM 126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101 ENKKYLVIVLEHVSQGLFPDYLVKGRLPPEARKFFROIYALDFCHSYISICHRLKPE 160
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127 ETEQTLVLETVSGGEVFDYLVHGMKEKRAKROIYSAVYLSKNIHDKAE 186
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
161 NLLDEKNRIADFGMASLQVDSLLETSCGSPHYACPEVIGKDYGRADMSGCVI 220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
187 NLLDDQMKIKIDFGSNTFSLCNKIDFGCSPYAPAEFSGKDYGPEDVWSLVI 246
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
221 LFLVGLALPDDDNLRQLLEKVRGVFHMHPHTIPDCOSLRGIEVEPEKRLSEIO 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 LYLTVSGSLPFDQNLKELEKRYLKGRIPEYMSDCENLKFLVLPQRSSLDIM 306
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
281 KHWYLVGKHEPD-PCLEPAPGRVAMRSLPSNGELDPVLESMAISGCR--DREEL 335
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
307 KQRMNVGYDEDELKPFIE-----PDKQIDEGRIELKIQ--POLGFKKAI 352
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
336 HRELRSSENOEKMTYLLDRKERYSCF---DQDLPRNDVPPRRKRVSPMSLRCK 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
353 LESVEKRFEDIHATYLLGERKSDMASEITMAOSLISHSIN-----SSLSGHPA 406
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
393 RPERKSEMEVLTIDAGGSGPVPTRRALEMAQHSQSRSVSGASTGLS--SPLSPR 449
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
407 CVLTRE--HTSSASAGSSASP-----SRYSRSATATGASITGASALASANAQ 455
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
450 SPVFSFSPGAGDEANGGSPSTKTO---TLPSKPGGCGAGDEPPPSANST-PLPGR 505
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
456 HQSSAPSSGSSSSSSSSSSQDAATAAGTIVMSGTRHGVQMAOPTSRQATISLDP 515
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
506 PGSRSSGSGT-----PLHSLPHTPRAS---PGTPTGTPPPSPG-----GVGAA 548
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
516 PSYKPSNTTIOAIPIPLFNSTSTSSAOPSTIGTRKADKGRIPLNSTAVQGR 575
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
549 WRSRLNIRNSFLGSPRHRKRM---OVPTAEEMSSF--TPSSPELAKRSMGNTFS 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
576 TATGAVANNGGISHRHAQOOQYMNQUTSSTMKSLINKTPAAGTAAATSSSSSSAT 635
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
602 LDKEEQIFLVAKDKPLSKIKADYHAF-----STPSLSHVSLSQTS 643
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
636 -----STAPLQKSGSOISHAPTEPVIREDDENNENONGNVP-LIGGVGPQTS 683
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
644 FRAEYKASGSPVQKPRFOVDISSSEGPSPRRDGGGIGIVFTLISGP----- 698
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
684 PAVQVPEPDATSSDKEQOQK--ASSETPKR-----SNPTVMQN 721
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
699 ---SRFRVETIOAQLSTHDPSPVQALADEKNGAQTTPAGAP-----RSI 744
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
722 LHLNLSLKSLDSSAATSYETPRRGIA-----GRSEPSAATPRRRHQTIVVDARHL 774
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

OY 745 QPPGRRPP 753  
 DB 775 QTPPTDRP 783

## RESULT 10

148609  
 probable serine/threonine-specific protein kinase (EC 2.7.1.-) kem - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 04-Mar-2000  
 C:Accession: I48609; S31333  
 R:Ingls, J.D.; Lee, M.; Hill, R.E.  
 Mamm. Genome 4, 401-403, 1993  
 A:Title: Emk, a protein kinase with homologs in yeast maps to mouse chromosome 19.  
 A:Reference number: I48609; MIM:93364122; PMID:8358177  
 A:Accession: I48609  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-774 <ING1>  
 A:Cross-references: EMBL:X70764; NID:957919; PIDN:CAA50040.1; PID:957920  
 R:Ingls, J.D.; Lee, M.; Hill, R.E.  
 submitted to the EMBL Data Library, January 1993  
 A:Description: A novel protein kinase with homologues in yeast maps to mouse chromosome 19.  
 A:Reference number: S31333  
 A:Accession: S31333  
 A:Molecule type: mRNA  
 A:Residues: 1-698; 'K', 700-702, 'GRGLPTAAKASCTSGTCYAGAHQATRT', 731-774 <ING2>  
 A:Cross-references: EMBL:X70764  
 C:Genetics:  
 A:Gene: emk  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom  
 C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
 F:51-304/Domain: protein kinase homology <KIN>  
 F:59-67/Region: protein kinase ATP-binding motif

Query Match 19.2%; Score 788; DB 2; Length 774;  
 Best Local Similarity 30.1%; Pred. No. 2.8e-21;  
 Matches 238; Conservative 119; Mismatches 261; Indels 172; Gaps 27;

```

30 YVGPRLKLTGKQGLVAVLVCITGOKVAIKIVNREKISEVLMKVERIALIKLIE 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
49 HIGNRYLKTIGKGNFVKVILKRIHILGKEVAVIIDKTOINSSLOKLFREVIMKVLN 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
90 HPRVILKHDYVENKRYLVLEHVSQGLFPDYLVKGRLPPEARKFFROIYALDFCHS 149
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109 HPRVILKHDYVENKRYLVLEHVSQGLFPDYLVKGRLPPEARKFFROIYALDFCHS 168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
150 YSICHRDLKPELNLDEKNRIADFGMASLQVDSLLETSCGSPHYACPEVIGKDYD 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
169 KFLVHNDLKAENLLDADMKIKIDFGSNTFSLCNKIDFGCSPYAPAEFSGKDYGP 228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210 RADMSGCVIYLVGLALPDDDNLRQLLEKVRGVFHMHPHTIPDCOSLRGIEVE 269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
229 PEVDVMSLVGLVLYLVGSLPFDQNLKELEKRYLKGRIPEYMSDCENLKFLVLP 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
270 PEKRLSEIOKHWPYLVGKHEPDPCLEPAPGRVAMRSLPSNGELDPVLESMAISG 329
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
289 PSKRGTIEQIMKDRMNVG--HED-----ELKRYV-EPILITGP- 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
330 RDR-----ERLRELSEENOEKMTYLLDRKERYSCFDLPSPCEDLPDPRNDV--- 376
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
326 RDRVDGVNGLHTEIQSLVGQRYN-EVMATYLLGLGKSEPEGDTILKPRPSADLTNS 384
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
377 -----PRKRVDS-----PLSRHGK-----RRPEKSMENVLSI 405
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
385 SAPPSPKRVORSVANRKRSSDOAVPAITPNSYKRTGSNNAENRPEETGRKASS 444
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
406 TDAGGSGSPV--TRRLLEMAQHSQSRSVSGASTGLS--SPLSPRSPVFSFSPGP-- 460
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
445 T-AKVPAFPLGLDKRKTTPA--PSTNSVLSSTINRNRNPLLD-RASLGASATQNGND 499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
461 ---AGDEARCGSPSTKQTLPRGPRGCGAGDEPP-----PESANSTPLPDPGSPRS 511
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 500 SLTPGSRSTASAAVSAARPHQKMSASVHPNKAISGLPTESNCEVPRPSTAPQR 559  
 QY 512 SGGTPLHPLTPRASPTGCTT-----ppspg 541  
 Db 560 --VPVAPSAHNISSSGADPTNTPROVSSRSTFHAGOLQVQDQMLPYGVTPASPS 616  
 QY 542 GGVGAARSLNLSIRNFTLSPFRHRRMAYTAEE-MSILTP-----584  
 Db 617 GHSGR--KGASGLTFSF--TSKFRRLNLPESKDRVETLRPHVYVSGGCTDKDEFR 672  
 QY 585 ESSBELAKRSW-FGNFISLDKEEQIFVLKDKPLSSIKADIYHAFSLIPSLSHVLSOTS 643  
 Db 673 EAKPRSLFTWSMTKSTSMENEMREIRKVLADANSQSELHERVLL--CVHGTGPHEN 730  
 QY 644 FRAEYKASGGSVFOKPYRFOVDISSSEPPSPRDGSGGGTIVTTLISGSPRRK 703  
 Db 731 F-----VQEMEVCK-----LPLRLSLNG-----VRFKRISGTSMAFK 762  
 QY 704 RVEITQAOI 713  
 Db 763 NIASKIANEL 772

## RESULT 11

gik protein - chicken  
 N/Alternate names: Qln-induced kinase  
 C/Species: Gallus gallus (chicken)  
 C/Date: 17-Nov-2000 #sequence\_revision 17-Nov-2000 #text\_change 17-Nov-2000  
 C/Accession: J07500  
 R/Xia, Y.; Zhang, Z.; Kruse, U.; Vogt, P.K.; Li, J.  
 Biochem. Biophys. Res. Commun. 276:564-570, 2000  
 A/Title: The new serine-threonine kinase, Qlk, is a target of the qln oncogene.  
 A/Reference number: J07500  
 A/Contents: Embryo fibroblasts  
 A/Accession: J07500  
 A/Molecule type: mRNA  
 A/Residues: 1-798 <XIA>  
 A/Cross-references: GB:AF219232  
 C/Comment: This protein, a member of the AMPK/SNFI family of serine/threonine kinases,  
 C/Genetics:  
 A/Gene: qik  
 C/Keywords: protein kinase

Query Match 19.0%; Score 780.5; DB 2; Length 798;  
 Best Local Similarity 29.6%; Pred. No. 5.3e-21;  
 Matches 244; Conservative 105; Mismatches 265; Indels 209; Gaps 25;

QY 17 LPHPHPPPOAAOYGPYRLKTKGQGLVYKLGVCITGKVAIKIYNREKLSVLM 76  
 Db 10 VPAPSAOPRPLR-VGYDIERLTGKGNFAVVKLARHRTKQVAIKITDKTLDPSNLE 68  
 QY 77 KVERRIALIKLIEHPVHLKLDVYENKRYLYLVLEHVSGLGELDYLVKGRILPKARKK 136  
 Db 69 KIVREVOIKMLKLNPHIKLYIKLYQVEMTKDMLYVTEFAKNGEMFDHLSNGHSESARKK 128  
 QY 137 FQIYVSAIDFCHSYSTICRDLKPEMLLDKNNIRIADGMSLQVDSILETSCSPHY 196  
 Db 129 FQIILSAEYCHSHIIVARDIKTEMLLDANNIKLADGFGFNFYKSGEPLSTWCSPPY 188  
 QY 197 ACPEYIKKEDYGRADMSGVILFALLVGLPDDDLROLEKVGGVFMPPIPP 256  
 Db 189 AAPEYBEKEKEGPHLDWISGLVLYVCGSLPFDGPHLPILKQVNLGSRFPIPFMSR 248  
 QY 257 DCQSLKGMIEVEPEKRLSDIOIKHWPYLGKHPDPCLEBAPGRVAMRSLPSN-GEL 315  
 Db 249 DCEFLIRMLVVDPTKRITISQIKQHKW-----QADPSLRQOQSLSFMQNNNSMLGDY 303  
 QY 316 DPVLESASIGCFPRDRRLRLRLSEBENQKMTIYLLIDKERRPSCEDDLPFRNDY 375  
 Db 304 NEOVLGIMQTLGI--DRQRTVESLQNSSYNHFAAIYLLERLKEYS-----349  
 QY 376 DPKRKVDYPMLSIRGKRPRER-KSMETVLSITDAGGGSPVPTLRALMAQHSQSRSVS 424

Db 350 -----SOLSSRPATGRQQRPSSEISN-----AENPQSLISETL- 384  
 QY 435 GASTGLSSP-----ISSPRSPVFESPEPGADBARGGG-SPTSKTQTL 478  
 Db 385 -RSSILYQQPQSLQPSLQAEEMDCMNNPLQPVF-FPVDPNNGFLRNNSISPSLLETT 442  
 QY 479 PSRGPBGAGAGEPPPPARSTPLRPPGSPR--SSGTPPLHPLHTP-----RASPT 529  
 Db 443 ISEEVN--QEKLEDEIKAYDPIRIPNTSRHRLAEVTFHFYQAPPCIVISSASPT 500  
 QY 530 -GTPGTTPPS-----PGGVGAAMRSRLNSINSFLGSPFRHRRKQVPT 575  
 Db 501 EGTSSDSCITSSNDSSVALSSCLAQVWTSAPATAMTS--AFLAS-QSDAPVLQVGG 556  
 QY 576 AEEMSLTPESSPE-----LAKRSWFGNFISLDKEEQIFVLKDKP 616  
 Db 557 CMGASILPVSFQEGRRASDPTSLTGLKAFRQLKRNARAKFLGLNK-----604  
 QY 617 LSLIKADIVHAFSLIPSLSHVLSOTSFRAEYKASGGSVFOKPYRFOVDISSSEGPS 676  
 Db 605 -----IKGFARQVCGSSSRRAARSAM-----S 626  
 QY 677 PRDGGSGGITYSVTFLISGPRRFRKRYETI--QAQLSTHDQPSYQALADEKNGA 733  
 Db 627 PFQHAQPTCTIYS-----SSGSSREGRLNLEEVLOQQRMLQLOHHO-----667  
 QY 734 TRPAGAPRSLQPPGPRDPDELSSSPRGRPPKDKILATNGTP 776  
 Db 668 -----LLQF-----ACPOTSQTSATNGLP 686

## RESULT 12

hypothetical protein 22E5.8 - fruit fly (Drosophila melanogaster)  
 C/Species: Drosophila melanogaster  
 C/Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
 C/Accession: T113741  
 R/Murphy, L.; Harris, D.; Barrell, B.  
 Submitted to the EMBL Data Library, April 1999  
 A/Description: Sequencing the distal X chromosome of Drosophila melanogaster.  
 A/Reference number: Z17668  
 A/Accession: T113741  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-1398 <MUR>  
 A/Cross-references: EMBL:AL031765; NID:e1371523; PID:e1329905; PIDN:CAA21125.1  
 A/Intons: 205/3; 227/1; 322/3; 688/3; 782/3; 814/2; 1363/3  
 A/Note: EG:22E5.8

Query Match 18.7%; Score 767; DB 2; Length 1398;  
 Best Local Similarity 28.2%; Pred. No. 2.5e-20;  
 Matches 251; Conservative 112; Mismatches 324; Indels 202; Gaps 24;

QY 9 GGGSPAHPLPHPHPOHA-----QYGVPRLEKTLGKQGLVYKLG 51  
 Db 105 GGGST-----PGSPPTSAAVAGAGISGKDLTKKEPMVGVYDIERITGKGNFAVVKIA 158  
 QY 52 VHCITGKVAIKIYNREKLSVLMKYVEREIALIKLIEHPVHLKLDVYENKRYLYLVE 111  
 Db 159 RHRTKMEVAIKITIDKSDLODTNLOKVRREVMIRKLPHIKLYQVEMTKMITYISE 218  
 QY 112 HVSGLGELDYLVKGRILPKARKFFROIVSALDCHSYSTICRDLKPEMLLDKNNIR 171  
 Db 219 YASGEIFDIYAKYGRMSASAARFKMOTIISAVECHKGIIYHRDLKAPNLLDLMNTRK 278  
 QY 172 IADFGMASLQVDSILETSCSPHYACEVYIKGEYKGRADMSGVILFALLVGLALPF 231  
 Db 279 IADGFSNHFRPGEILLATWCGSPPYAAAEVEFGKQYTGEIDWISGLVLYVCGALPF 338  
 QY 232 DDDNLRLQLEKVKKGVFMPPIPPDCQSLKGMIEVEPEKRLSDIOIKHWPYLGKHE 291

```

Db 339 DGSITQSLRDLRVLSGRFRIPFPMSECEHLIRMLVLEPTRYTIDQIKRHHMM----- 392
QY 292 PDCLEAPRGRVAMRSL-----PENGELDPVLESMA-SJGFRDRERHRLRSEENO 346
Db 393 ---CPLEHLLVLAQYVLAERQTSVEPSEDLIRMAEYVIGSDKTR--ASLKNNTYDH 447
QY 347 EKMIVYLLDR---KER-----YSCEDODLPRNDVPPRKR-----V 382
Db 448 VAAIYLLQDRVSHKQKQSNGLGASALASSTASRMITSNRDHPQTQOQSOQOSTIST 507
QY 383 DSPMLSHKRRPRKRSMEVLSITDAGGGSFV-----PTRLAEMAQHSOR 429
Db 508 SSIILADQCKRSLRSHQVLMSERNAHAGATVPDPGPYAKYGLQLPRLTGHSHL 567
QY 430 SRSVSGASTGSSPLSPRSPVSESPGAGDEARGGSPSK---TQTLPSRGR-- 484
Db 568 TGYLNGGVEVDASGIFLP---MRYTLPPLAASPAPSCSTSRVGHSLSSSPRSH 623
QY 485 ---GGAGEQPPPPSARSTPLPG----- 504
Db 624 RPAVLSLIDNPNPLANLRCREMMAGG-----PGAVGVPPLASKQOLHQTISEIILQ 677
QY 505 -----PGSPNSSGGTPLHSPHTPPASPTGTPPTPPSPGGG 543
Db 678 STEDCRALLQOSTAVAEQKDPKPAESSVGVPP-----PPASTPTSTAGPES--- 726
QY 544 VGGAMRSRLNSINSPFSRFRHKKQVPTAEEMSLTPESSPELAKRSFNGFNISID 603
Db 727 -GAPCPGEING-----KTIKMSSSSSSFDSTANLQGSF--RKMSA 765
QY 604 KEQDIFLVLDKPLSLIKADIVHAFSLPSLSHVLSTQSFRAEKASGGSPVFOKPVRF 663
Db 766 EASLFTGLDESLPVEQRTKRHVHVGSTNGSGDSGGETNDAK---SNGDSRSEKVLVA 822
QY 664 QVDISSEGEPPSPRRD-----GSGGGIYSVTFTLLISGPRRKRKYVETIOQLL 714
Db 823 QGSSSTEGCCTTQDGNPDSASQESKSGNGSGNANGPSSHSSDILRLVGTASG-- 880
QY 715 STHDPQSVQALADEKNGQTRAPGAPRSLQPPGPRPDELSSPPRRP 763
Db 881 QSHMRS-YASSSSSGVLAGSAGSYKSLSONLSRGS---SKSMCSGP 925

```

## RESULT 13

T07415  
probable serine/threonine-specific protein kinase (EC 2.7.1.1) PKIN1 - potato  
N/Alternate names: SNF1-related protein kinase  
C/Species: Solanum tuberosum (potato)  
C/Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 20-Jun-2000  
C/Accession: T07415  
R:Halford, N.G.  
submitted to the EMBL Data Library, March 1996  
A/Reference number: Z16022  
A/Accession: T07415  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-504 <HML>  
A/Cross-references: EMBL:X95997; NID:91216279; PIDN:CAA65244.1  
C/Genetics:  
A/Gene: PKIN1  
A/Intons: 62/1; 123/3; 184/3; 228/3; 290/3; 320/3; 348/3; 389/3; 466/3  
C/Function:  
A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin  
C/Superfamily: AMP-activated protein kinase; protein kinase homology  
C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
F;15-269/Domain: protein kinase homology <KIN>

Query Match 18.7%; Score 766.5; DB 2; Length 504;  
Best Local Similarity 41.1%; Pred. No. 1.1e-20;  
Matches 157; Conservative 74; Mismatches 124; Indels 27; Gaps 6;  
QY 30 YGPRLEKTLGKQGTGLVGLVHCITGQKVAIKIVNREKL-SESVLAKVEREIALIKLI 88

```

Db 13 YLKNRYVKGTLGHQSFQKVAIEHLNIGHKVAIKILNRRKKKTPDMEEKLRREIKICLF 72
QY 89 EHHVYVLAHLYENKKTLYLVLEVSGSELDYLVKKGRLLPKARKFRQIVSLDCH 148
Db 73 VHHVIRLYEVLETPDIIVMEYVKSSELDYLVKGRLEQARKEADKATFOQIIGVEYCH 132
QY 149 SYSICHRDLPENLIDDEKNNIRIADFGMASLQVDSILLETSCGSPHYACREVIKGEYD 208
Db 133 KNAVYHDLKPENLIDARRKVKIADFGIMDGHFLTKSCGSPHYAAAEVYSKITYA 192
QY 209 GRADWMSGVILFALLVGLPFDNDNLKQLLEKVRGVFHPHTIPDCSLLRGMIEV 268
Db 193 GPEVDWMSGVIIYALLCGTLPEDDENIPLNFKKIKSGVYTLPSHLPLARDLIPRLIV 252
QY 269 EPEKRSLIEDIQHWPVLGKHEPDCLEPARGRVAMRSLPENGELDPVLESMA-SJG 328
Db 253 DPKRKISVPIRQHQWF--KILPRYLAVPPDARQHLK-----KLDEETLQVSRMGL 304
QY 329 FRDRERLHRLRSEENOEMKIYVLLDRKERYPS-----CEDODLPRNDVD 376
Db 305 --DRDQLDLSLQRIQDQDAVAYVLLVDRMSASSGYLGAEFQESVDCYSGILFPNDL- 361
QY 377 PPKRVDSPMLSRHKKRPERK 398
Db 362 ---QLSTGNGVSESLRRPFRK 380

```

## RESULT 14

A56009  
serine/threonine-specific protein kinase (EC 2.7.1.1) NPK5 - common tobacco  
C/Species: Nicotiana tabacum (common tobacco)  
C/Date: 28-Apr-1995 #sequence\_revision 28-Apr-1995 #text\_change 16-Jun-2000  
C/Accession: A56009  
R:Murataka, T.; Banno, H.; Machida, Y.  
Mol. Cell. Biol. 14, 2958-2965, 1994  
A/Title: Characterization of tobacco protein kinase NPK5, a homolog of Saccharomyces  
tase of Saccharomyces cerevisiae.  
A/Reference number: A56009; NUID:94217693; PMID:8164654  
A/Accession: A56009  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-511 <MUR>  
A/Cross-references: GB:D26602; NID:9496384; PIDN:BA05649.1; PID:9496385  
C/Function:  
A/Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threo  
C/Superfamily: AMP-activated protein kinase; protein kinase homology  
C/Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein  
F;17-371/Domain: protein kinase homology <KIN>  
F;25-33/Region: protein kinase ATP-binding motif  
F;48-67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted  
F;147,151/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 18.7%; Score 766.5; DB 1; Length 511;  
Best Local Similarity 44.4%; Pred. No. 1.1e-20;  
Matches 151; Conservative 67; Mismatches 107; Indels 15; Gaps 5;  
QY 34 YLEKTLGKQGTGLVGLVHCITGQKVAIKIVNREKLSSEVL-MKVEREIALIKIEHPH 92  
Db 19 YKLGKTLGIGSPGKVAIEHTLTGHKVAVKILNRRKIKNMMEKREVRREIKILRLEMPH 78  
QY 93 VKLHLYENKKTLYLVLEVSGSELDYLVKKGRLLPKARKFRQIVSALDCHSYSI 152  
Db 79 IILLYEVEIPSDIIVMEYVKSSELDYLVKGRLEQARKEADKATFOQIISVEYCHRMV 138  
QY 153 CHRDLPENLLDEKNNIRIADFGMASLQVDSILLETSCGSPHYACREVIKGEYDGRA 212  
Db 139 VHRDLKPENLIDSKMVKVKTADFGLSNIMRDLKTSKSGSPHYAAPAEVIGKLYAGPEV 198  
QY 213 DMWMSGVILFALLVGLPFDNDNLKQLLEKVRGVFHPHTIPDCSLLRGMIEVEPEK 272  
Db 199 DWSGVILYALLCGTLPEDDENIPLNFKKIKGMSLPSHLISAGARDLIPRLIVDPK 258



